

# Platform to Support Video Encoding

● Tatsushi Otsuka

Since the days of analog TV broadcasting, Fujitsu has been engaged in the development of LSIs for compressing video images compatible with MPEG-2. In recent BS/terrestrial digital broadcasting, the broadcasts themselves are data compressed with MPEG-2, but more efficient recording can be achieved by reconvertng them into H.264, which features even higher compression efficiency than MPEG-2. In this way, the fields in which LSIs for compressing video images can be applied are expanding. These video image compression processes, including audio processing and the processing of data streams generated by compression, need to match the various types of data to be processed. In addition, video image compression, which characteristically uses a massive amount of memory and handles a large amount of data, essentially requires an external DRAM connection. Accordingly, another important factor in such compression is the configuration of a memory system. In relation to these development requirements, this paper presents the LSIs that Fujitsu has developed from the perspective of a platform. The building of a platform has allowed Fujitsu to accumulate design assets, the surrounding environment and know-how, and allowed it to efficiently develop new LSIs.

## 1. Introduction

International standards of digitization of video images for compression and decompression include MPEG-2 and H.264/AVC.

In a narrow sense, video image compression processing is compression of images. However, since the days of analog TV broadcasting, Fujitsu has been developing before others LSIs capable of not only compressing video images but also audio compression and processing the compressed data into a stream according to the application, all in one chip. These LSIs, which are called encoders because they encode multimedia data including video images and audio into data compatible with the respective standards, have been used for applications such as hard disk recorders and video recording boards for PCs.

Along with the subsequent standardization of H.264/AVC as a more efficient compression scheme, we developed encoder LSIs compatible with it, and they have been used for applications such as video transmission and video cameras. In addition, as the broadcasts themselves have come to be offered as

data compressed with MPEG-2 in BS/terrestrial digital broadcasting, we have also developed LSIs for recompression into H.264/AVC, which has higher compression efficiency; hence they are called transcoders. Our LSIs provided significantly longer recording times than recording as MPEG-2 data, and combining this with a greater hard disk capacity we have brought into existence applied products that are capable of recording all the broadcasts of multiple channels for a period that ranges from a few days to a few weeks.

This paper describes the characteristics of the respective types of data and presents the LSIs we have developed from the perspective of a platform.

## 2. Characteristics of respective types of data

### 2.1 Video image data

Video image data are composed of about 30 frames of still image data per second (for NTSC areas including Japan; 25 frames per second for PAL areas such as Europe). Video images of standard-definition television (SDTV) in the days of analog TV broadcasting

include a data volume of approximately 160 Mb/s and those of high-definition television (HDTV) approximately 1 Gb/s. With MPEG-2 and H.264/AVC, these data are efficiently compressed by making use of correlation between images, and this has successfully compressed them into data of a few Mb/s to a few tens of Mb/s.

The reason why such high compression rates can be achieved is that there are inherent high correlations between each of the still images that constitute a video. Significant data compression is made possible by efficiently extracting the differences between each still image and making only those portions into data.

To make use of the correlations between images, however, data of multiple images must be stored in memory (called frame memory) before processing. Having a larger number of frames that can be stored in the frame memory allows more efficient compression, which means that video image compression requires a large amount of memory.

When only one still image needs to be processed, the input image can be once stored in the frame memory and sufficient time can be spent on the necessary processing for writing to the recording medium. With video images, however, input to the frame memory must continue and, at the same time, data processing between images that have been input to the frame memory must be performed without delay.

In this way, video image processing is characterized by handling large amounts of memory and data. For integration into LSIs, this is generally achieved by dedicated hardware.

## **2.2 Audio data**

Audio data are data resulting from sampling sound, which is an analog quantity, at a certain time interval. When the two left and right stereo channels are respectively sampled at 48 kHz with 16 bits/sample, the data volume is approximately 1.5 Mb/s. With compression schemes such as AC3 and AAC, these audio data are compressed into data volumes of a few hundred kb/s.

In terms of data volume and compression rate, audio data are incomparably smaller than video image data and so they tend to be neglected. However, audio processing involves multifold arithmetic processing of data of a few hundred to a few thousand samples, which is not easy for an embedded system CPU.

In addition, audio processing systems themselves are developed on the assumption of digital signal processor (DSP) processing of a few hundred MIPS and there is a tendency for new systems to be developed relatively frequently.

For this reason, DSPs and high-performance CPUs are generally used together with signal processing assist circuits.

## **2.3 Stream**

As opposed to original data of video images and audio, data encoded by various schemes of compression processing are specifically referred to as video elementary streams and audio elementary streams. These data are subjected to header processing and appended with various types of management information so that they can be made into data adapted to applications, which are called system streams or simply streams.

A data format called MPEG2-PS (program stream) is used for DVDs and a data format called MPEG2-TS (transport stream) for digital broadcasting including terrestrial, BS and CS broadcasting and Blu-ray disks. Other data formats may also be used, such as MP4 and MOV, according to the application.

In this way, data processing is applied according to the application, and this is stream processing.

While stream processing handles data of a few Mb/s to a few tens of Mb/s, for video and audio elementary streams, which account for most types of data, processing mostly includes addition of headers and division and integration in certain sizes and the data contained in elementary streams are passed on almost as they are. For this reason, a configuration combining dedicated hardware or a CPU with an assist hardware circuit is generally used for stream processing.

The assist hardware in this case is not positioned as a way to make up for the insufficiency of the CPU capacity such as an assist circuit in audio signal processing. Rather, it is regarded to act in place of the CPU and perform the routine processes that cannot be efficiently performed by sequential processing with CPU programs, such as the direct memory access (DMA) feature.

### 3. Methods of memory system realization

What is as important as processing various types of data is how to realize a memory system. The methods of achieving a memory system can be classified into the following two types.

#### 1) Independent memory

This type of method provides dedicated memory for each application of processing such as video images, audio and streams. The design is optimized for each type of processing and the design of the respective type is simplified but physical restrictions apply. In particular, video image processing requires memory for storing several frames' worth of image data because correlations between images are used, which makes DRAM external to the chip essential.

For audio processing and stream processing, providing SRAM of the required capacity in the chip allows the realization of processing with independent memory.

#### 2) Shared memory

This type of method uses DRAM connected outside the chip and it is shared by the respective processing blocks for different applications corresponding to certain units of addresses, rather than providing dedicated memory for each application of processing.

The respective processing blocks use the DRAM, which prevents the individual processing blocks from freely accessing the necessary data. For this reason,

certain measures must be taken in the circuit such as making processing wait, or providing small-capacity SRAM on the chip as buffer memory and estimating when the respective processes require data to read ahead when available in order to prepare data.

### 4. Configurations of respective LSIs

Of the LSIs that have been developed by Fujitsu, **Figure 1** shows a classification of those relating to video image compression processing from the perspective of the platforms presented below.

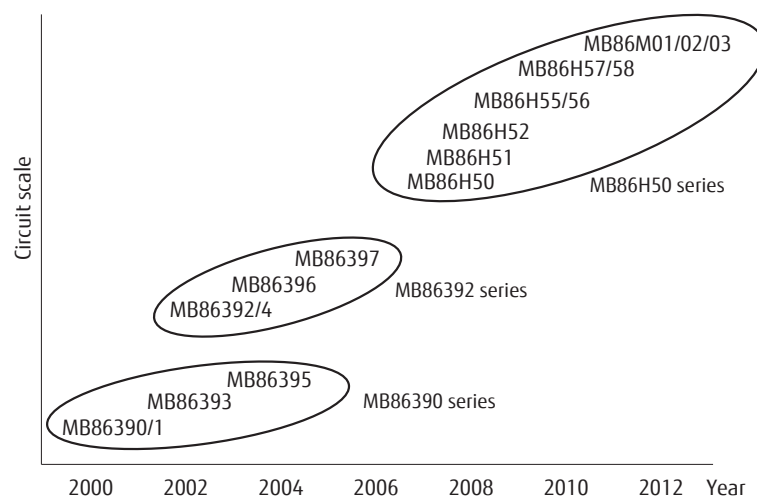
#### 4.1 MB86390 series

MB86390<sup>1)</sup> has the core of the already-developed encoder LSI exclusively for MPEG-2 video as the video image processing block and integrates DSP for audio processing (HiPerion1), dedicated hardware for stream processing and CPU (SPARClite) for overall control of the LSI.

A schematic block configuration is shown in **Figure 2**.

The video input data are directly led to the video image processing block and audio input data to the audio processing block and the elementary streams resulting from the respective processing are connected to the stream processing block.

The configuration with totally independent memory has audio and stream processing blocks integrating SRAM required for the respective type of processing,



**Figure 1**  
Classification of LSIs from perspective of platforms.

and a video image processing block integrating a DRAM controller, and it is provided with external SDRAM dedicated to video image processing.

## 4.2 MB86392 series

MB86392<sup>2)</sup> is the first LSI developed in view of a platform.

A schematic block configuration is shown in Figure 3. MB86392 is a full-duplex codec capable of simultaneous encoding and decoding but the figure is an excerpt showing the encoding block only.

The MB86390 series was characterized by having an optimum design for the respective processing blocks and compactness. However, the close coupling

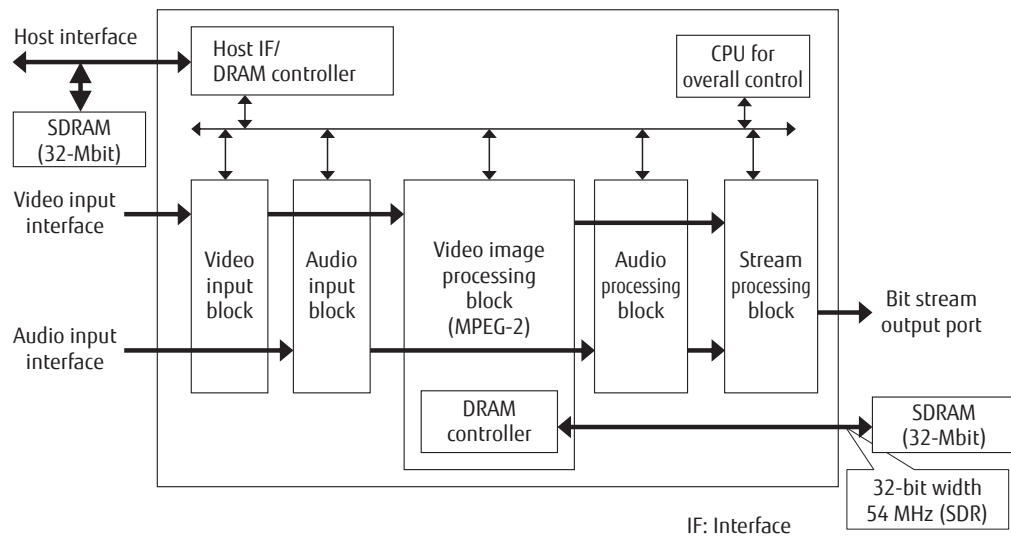


Figure 2  
MB86390 schematic block configuration.

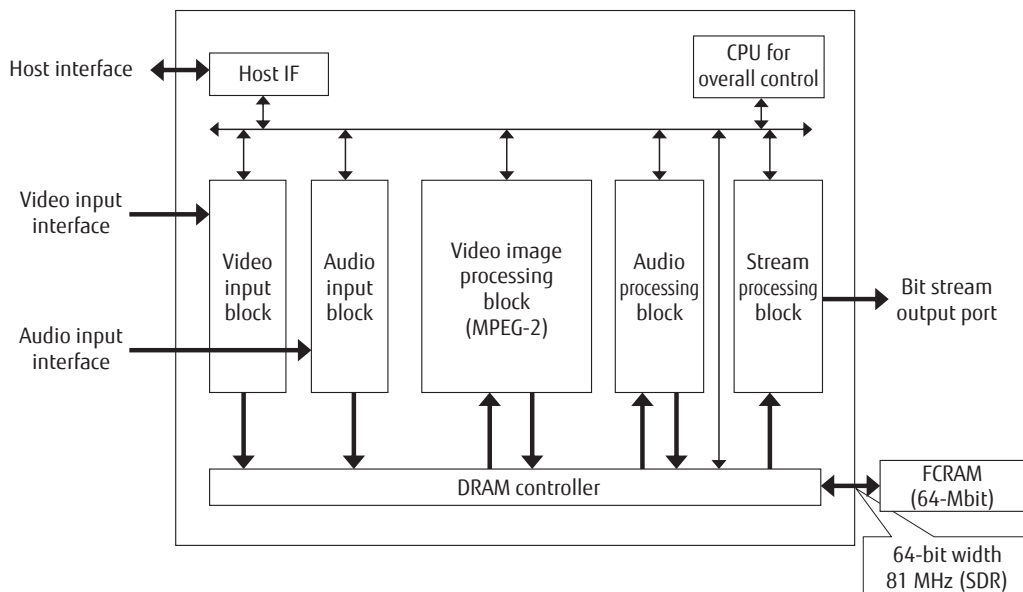


Figure 3  
MB86392 schematic block configuration.

between the processing blocks made it difficult to extend the products such as by adding functions.

To deal with this problem, MB86392 is configured based on the following principles with the focus on extensibility and flexibility.

- The processing blocks are laid out flat on the DRAM controller.
- Couplings between the processing blocks are eliminated whenever possible.
- Input/output data of the processing blocks pass through the buffer in the DRAM.
- The buffer addresses of the data in the DRAM are controlled by the CPU.

These contrivances have improved the independency of the respective processing blocks and made it easier for units of processing blocks to be replaced and functions to be added.

Meanwhile, the input/output data of the processing blocks always pass through the external DRAM to share the memory. This is disadvantageous to video image processing, which requires large volumes of data. Specifically, even when DRAM access is required for video image processing, other types of processing may be using the DRAM to prohibit the access, which causes idling in the video image processing. If this occurs frequently, the actual video processing may not proceed in time even if the DRAM capacity is sufficient

in terms of the bandwidth.

To avoid this phenomenon, each of the circuits that require data input in the respective processing blocks is provided with small-capacity SRAM as buffer memory to read ahead for preparing data before the processing requires data. Small-capacity SRAM is also provided as buffer memory for the output to prevent the circuit from being idle when waiting for output. These SRAM buffers also function to bundle fractional accesses into a burst access to the DRAM.

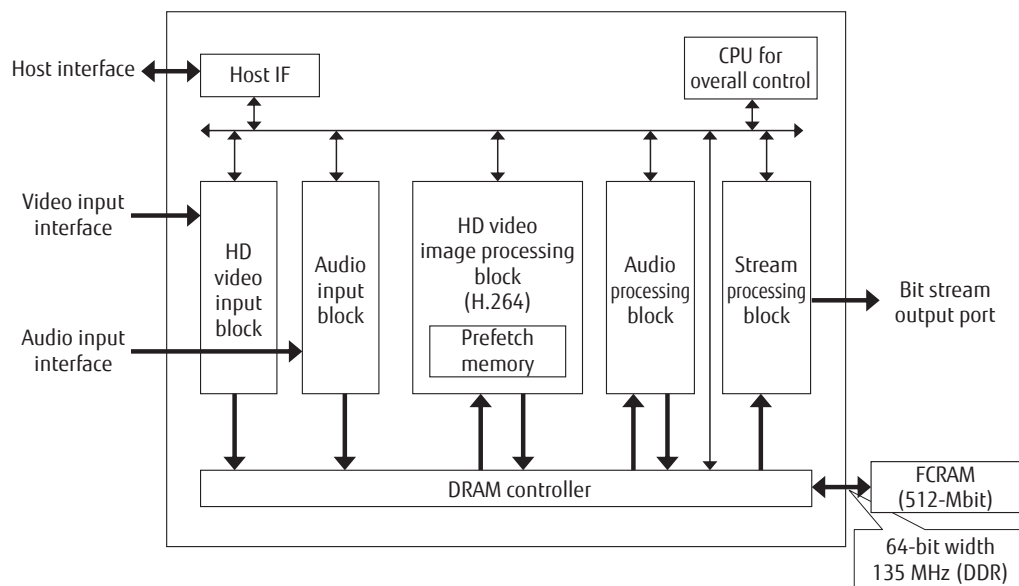
### 4.3 MB86H50 series

The MB86390 and MB86392 series were LSIs compatible with the MPEG-2 standard for analog SDTV.

In contrast, MB86H50<sup>3)</sup> provides a new platform compatible with HDTV and the H.264/AVC standard.

As shown in the schematic block configuration of **Figure 4**, the framework of the platform takes over the MB86392 series.

However, the volumes of video image processing have significantly increased with the HD screen sizes and the small-capacity SRAM for the respective processing block has become insufficient to absorb the effect of idling caused by the time to wait by arbitration of DRAM access. For this reason, the MB86H50 series is equipped with prefetch memory<sup>4)</sup> by means of large-capacity SRAM in the video image processing block.



**Figure 4**  
MB86H50 schematic block configuration.

To compare this to a CPU, the small-capacity SRAM of the respective processing block serves as the primary cache and the prefetch memory by means of large-capacity SRAM as the secondary cache.

#### 4.4 MB86M01/02/03

This subsection presents the latest LSI on the MB86H50 series platform. MB86M01/02/03 are transcoder LSIs that decode streams encoded by MPEG-2 or H.264/AVC, change the image size or compression rate and re-encode by MPEG-2 or H.264/AVC (see "H.264 Transcoder LSI Chip" contained in this magazine for the details). **Figure 5** shows a schematic block configuration from the perspective of a platform.

MB86M01 has an encoder and decoder for MPEG-2 and H.264/AVC respectively, and is equipped with multiple audio processing blocks and stream processing blocks. While its functions have been significantly enhanced from MB86H50, the MB86H50 series platform is used as the basis for its development and the assets developed with MB86392 are used for the MPEG-2 video image processing block.

### 5. Effects of platform

There are many benefits of a platform. In particular, we consider the following effects as significant based on our experience with the development of these LSIs.

#### 1) Accumulation of design assets

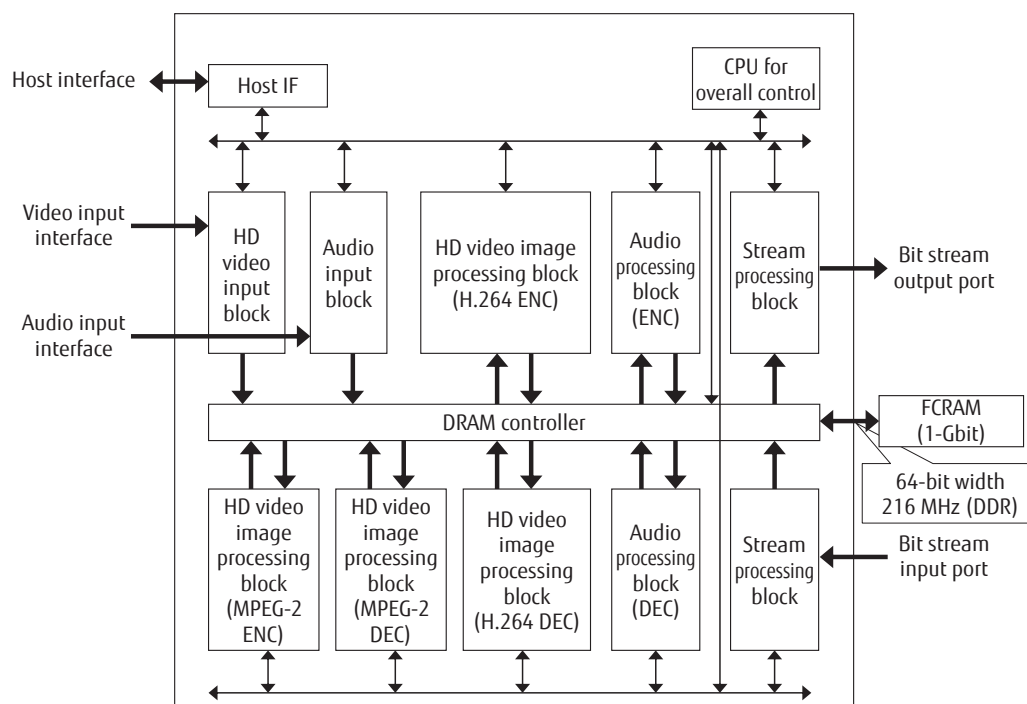
Design assets such as hardware and control firmware of the processing blocks are accumulated, which allows an engineer to concentrate on the new development portion for subsequent LSI development.

#### 2) Accumulation of know-how

There are various types of know-how. For example, feeding back the results of making practical use of the developed LSIs, such as estimates of the effective performance against the memory bandwidth and relation between the circuit scale and power consumption, allows for more accurate estimates for the subsequent LSI development.

#### 3) Improvement of surrounding environment

The building of a platform makes it easier to improve the surrounding environment such as the verification environment including simulators and emulators, various tools and evaluation environment



**Figure 5**  
MB86M01/02/03 schematic block configuration.

of the actual product and the design quality can be further improved for the subsequent LSI development.

As described in 1) to 3) above, the many assets accumulated and developed based on a platform brings higher efficiency to the development of new LSIs.

## 6. Conclusion

This paper has described the characteristics of video image, audio and stream processing and memory systems essential to media processing and presented from the perspective of a platform video image compression processing LSIs that have been developed by Fujitsu.

From now on, we intend to extend the platform to cover the 4K2K ( $4096 \times 2160$  or  $3840 \times 2160$  pixels) and 8K4K ( $7680 \times 4320$  pixels) image sizes that exceed

HDTV and achieve compatibility with High Efficiency Video Coding (HEVC), which is in the process of being standardized as a successor of H.264/AVC.

## References

- 1) K. Tanaka et al.: Single-chip MPEG-2 Audio/Video Encoder LSI. (in Japanese), *FUJITSU*, Vol. 53, No. 1, pp. 71–75 (2002).
- 2) K. Sakai et al.: VLSIs for Video Coding. (in Japanese), *FUJITSU*, Vol. 54, No. 1, pp.52–56 (2003).
- 3) H. Nakayama et al.: H.264/AVC HDTV Video Codec LSI. *Fujitsu Sci. Tech. J.*, Vol. 44, No. 3, pp. 351–358 (2008).
- 4) H. Matsumura et al.: A Low-Power H.264/AVC Codec with 1080p/60 Processing Capabilities. *Proc. IEEE Symp. Low-Power and High-Speed Chips (Cool Chips 09)*, IEEE CS Press, 2009, pp. 175–187.



**Tatsushi Otsuka**

*Fujitsu Semiconductor Ltd.*

Mr. Otsuka is currently engaged in development of video encoding LSIs.