

# H.264/AVC HDTV Video Codec LSI

● Hiroshi Nakayama   ● Yasuhiro Watanabe   ● Akihiro Higashi

*(Manuscript received December 4, 2007)*

Fujitsu has developed a video codec LSI for the compression/decompression of high-definition television (HDTV) video in conformance with the H.264/AVC standard, which is expected to become the next-generation video compression standard. This LSI uses a “prefetch memory control technique” that reduces external memory access — an important issue in the compression of HDTV video — by 25 to 45% in terms of data transfer rate compared with conventional memory access techniques. It achieves superior picture quality and low power consumption by optimal application of video-encoding algorithms accumulated by Fujitsu over many years. The LSI is embedded with two of Fujitsu’s original low-power memory units (fast cycle random access memory [FCRAM]) in a single package under the product name MB86H50. The entire chip including memory consumes only about 600 mW during audio-visual (AV) encoding of HDTV video, achieving low power consumption at an industry-leading level. With the above features, the MB86H50 can record, play, and transmit HDTV video at high quality via a wide variety of equipment including portable AV devices, hard disk recorders, and home network appliances.

## 1. Introduction

High-definition television (HDTV) has become very familiar to consumers thanks to the dramatic spread of terrestrial digital broadcasting and flat panel TVs. Recent years have also seen the commercialization of portable audio-visual (AV) equipment such as HDTV video cameras and digital still cameras as an addition to stationary digital AV equipment for viewing digital broadcasts and DVDs. We are entering an era where HDTV video can be captured and enjoyed by the individual.

To handle the huge amount of data in HDTV video, it is generally necessary to compress that video data while maintaining high picture quality whether it be digital broadcasts having restricted frequency bandwidth, IP-based video transmission, or recording to storage media having limited capacity. The MPEG-2 international standard

is widely used as a video compression technology and is finding widespread use in DVDs and digital broadcasting. Recently, however, the new H.264/AVC international standard, which is capable of achieving a compression ratio more than twice that of MPEG-2, has been attracting attention as the next-generation compression technology succeeding MPEG-2.<sup>1)</sup> Fujitsu, meanwhile, has been working on the development of an H.264/AVC-compliant HDTV video codec (compressor/decompressor) LSI with the aim of providing compact and low-power digital AV equipment for recording and playing HDTV video. In November 2006, it announced the MB86H50 that combines, in one package, a video codec LSI supporting the H.264/AVC High Profile, Level 4.0 specification adopted in next-generation DVDs with two units of Fujitsu’s original low-power fast cycle random

access memory (FCRAM). This product embodies Fujitsu's video and audio processing technologies accumulated over many years and its 90-nm semiconductor processing technology. It will enable the high-quality recording, playback, and transmission of HDTV video on a wide variety of equipment including portable AV devices, hard disk recorders, and home network appliances.

This paper focuses on H.264/AVC video codec processing in support of HDTV. It first surveys the problems associated with implementing this video codec in LSI form and then describes key technologies developed by Fujitsu to overcome those problems. In the following sections, please note that FCRAM is defined as "external memory" with respect to the video codec LSI, although both are combined in the same package.

## 2. Problems in LSI implementation

Though it achieves a compression ratio more than twice that of MPEG-2, H.264/AVC compression technology requires a processing load more than ten times that of MPEG-2. The reason for this can be explained as follows. During the development of the H.264/AVC standard, it was expected that compression technologies used for past standards such as the removal of temporal and spatial redundancies and efficient variable length coding would provide a basis for the new standard. But achieving a higher compression ratio was expected to require advanced schemes and evaluation experiments, and these were actively discussed and proposed at international standardization organizations. In the end, it was decided that a large number of complex encoding modes would have to be defined. Consequently, to design an LSI that can extract maximum performance from the H.264/AVC standard while achieving high picture quality and low power consumption, it is necessary to develop an algorithm that can select the most appropriate encoding mode in accordance with video features

at the time of encoding and that can easily determine optimal encoding-mode parameters. It is also necessary to develop dedicated circuits that can efficiently assist with the application of that algorithm.<sup>2)</sup> It must also be kept in mind that such a video codec LSI would normally have to access external memory very frequently due to the nature of HDTV video data. Some means is therefore needed to reduce the amount of data transfer with external memory if the cost, power consumption, and generated heat of not just the LSI but of the entire device are to be considered.<sup>3)</sup> In the following sections, we first present the hardware configuration of the video codec LSI developed by Fujitsu and then describe technology for reducing data transfer with external memory. In particular, we explain:

- 1) a prefetch memory control technique for efficiently preloading picture data to be referenced with the aim of reducing the amount of data transfer with external memory during motion estimation in the encoding process and
- 2) a reference-block-grouping technique for efficiently reading in dispersed reference images in the decoding process.

## 3. Video codec LSI overview

A block diagram of the H.264/AVC-compliant HDTV video codec LSI is shown in **Figure 1**. The video codec core that executes compression/decompression processing consists of three main units: preprocessing, basic codec, and entropy. Each of these units is duplicated operating at 108 MHz in parallel.

The preprocessing block operates only at the time of encoding. It analyzes the input picture, calculates temporal and spatial statistical quantities that indicate pattern or motion features, and performs coarse motion-estimation and spatial-prediction processing and other processes. The basic codec block performs fine motion-estimation and spatial-prediction processing, spatial/frequency transformation conforming

to the H.264/AVC standard, quantization, and other encoding/decoding processes. The entropy block is the unit that performs variable length coding/decoding conforming to the H.264/AVC standard. The H.264/AVC standard defines a variable-length-coding system called context-based adaptive binary arithmetic coding (CABAC), which, though highly efficient, is very complex. By using special circuit designs and two-level parallel processing, the entropy block achieves an average throughput of 1 bit per cycle or more, resulting in processing performance (20 Mb/s) that is sufficiently high in practical terms for the compression/decompression of HDTV video.

In addition to the above three units, the video codec core provides a data transfer block that performs prefetch memory control (to be described below), prefetch memory, and a built-in central processing unit (CPU) that performs H.264/AVC syntax analysis and picture-quality

control. This CPU achieves high picture quality by using the temporal and spatial statistical quantities calculated by the preprocessing block and an original rate control algorithm that allocates an optimal amount of code to texture patterns important to the human visual system.

The video codec LSI also incorporates a memory controller for controlling two 256-Mbit FCRAM units, video input/output blocks, audio input/output blocks, an audio codec core for performing audio encoding/decoding, a system encoder and decoder for controlling input/output streams in transport stream (TS) format, and a system controller for performing overall control. **Table 1** lists the main specifications of Fujitsu's MB86H50 chip that integrates the above video codec LSI and two FCRAM units into a single package through the use of system-in-package (SiP) technology. A photograph of the MB86H50 chip is shown in **Figure 2**.

While not described in detail in this paper,

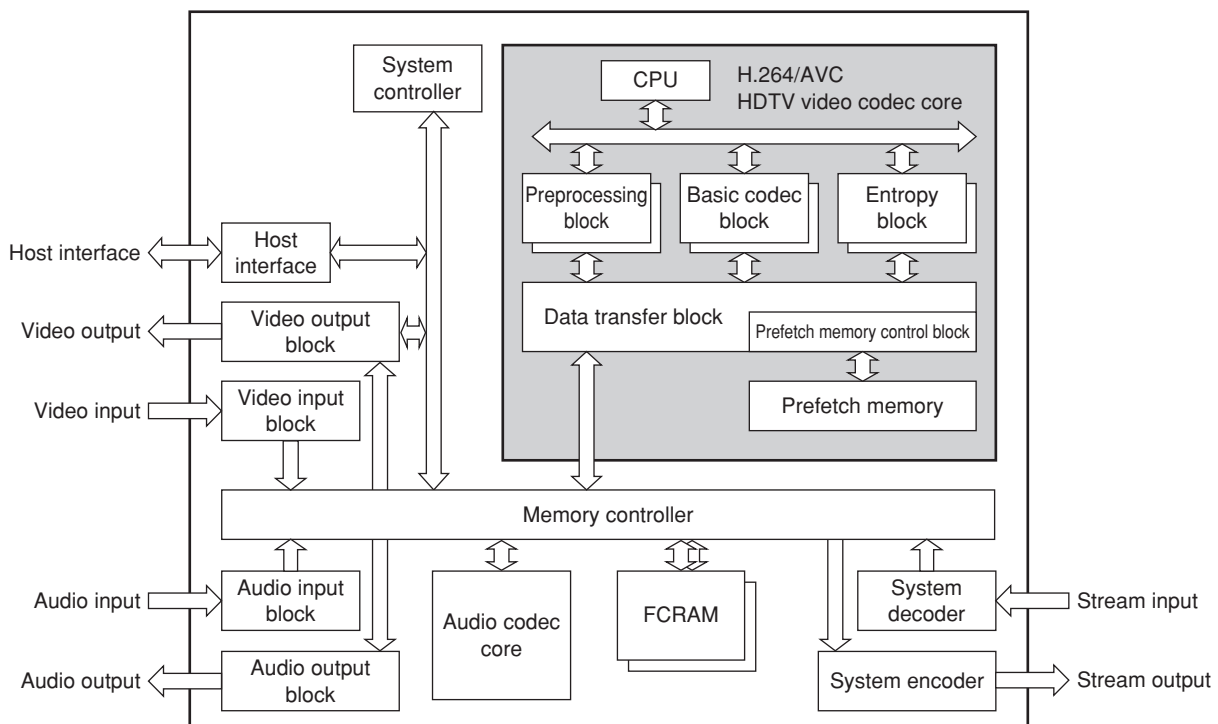


Figure 1  
H.264/AVC HDTV video codec LSI block diagram.

the audio codec core incorporated in this LSI also consists of Fujitsu technology developed and accumulated over many years. This technology conforms to various audio compression systems, as listed in Table 1, making the LSI applicable to diverse applications.

#### 4. Prefetch memory control

One of the basic principles of video compression processing in an MPEG system like H.264/AVC is to detect strong correlation between video frames, that is, regions having similar patterns, and to remove that redundant information. Detecting such regions of strong correlation between frames with high precision is therefore an important element of achieving high picture quality. This process, commonly called “motion estimation”, takes a macroblock, the basic unit of processing, and searches for a similar pattern in a temporally adjacent picture called a “reference image”. Here, performing an exhaustive search against macroblocks within a large region in the reference image will increase the precision of motion estimation, but a larger region means more accessing of reference images stored in external memory. In particular, in high-definition video like HDTV,

the magnitude of motion between adjacent frames is relatively large compared with the small images used in analog broadcasting, which means that motion-estimation processing will be accompanied by an extremely high occurrence of external memory access. The resulting demand on external-memory bandwidth and the increase in power consumption and generated heat all present major problems from a system viewpoint.

To solve these problems, the video codec core incorporates internal memory called “prefetch memory” and uses prefetch memory control to track motion in the vertical direction. This reduces external memory access and achieves high-precision motion estimation.

Based on the results of picture-quality evaluation simulations for different types of video, the optimal capacity of prefetch memory was decided such that two reference regions could be adequately prefetched for a picture frame to which forward prediction mode is applied (P frame) under the encoding of 1440 × 1080-pixel, 60-fields-per-second video.

The prefetch control scheme for vertical tracking determines whether a hit or miss has occurred in prefetch-area access for each instance of macroblock processing, and in the event of a miss, counts whether the miss is in the upper or lower direction. This count value is tabulated for each macroblock line and the resulting

Table 1  
MB86H50 main specifications.

Item	Specification
Video compression system	H.264/AVC High Profile, Level 4.0
Maximum picture size	1440 × 1080 @ 60 i 1280 × 720 @ 60 p
Maximum bit rate	20 Mb/s (max.)
System layer	MPEG-2 TS
Audio compression system	AC-3, Linear-PCM, MPEG-2 AAC, MPEG-1 Layer 2
Technology	90 nm CMOS
Power supply voltage	1.2 V (internal), 1.8 V (I/O)
Power consumption	600 mW (typ.) (for 1440 × 1080 i encoding)
Operating frequency	108 MHz (video codec) 135 MHz (external memory I/F)
Package	FBGA-650, 15-mm square

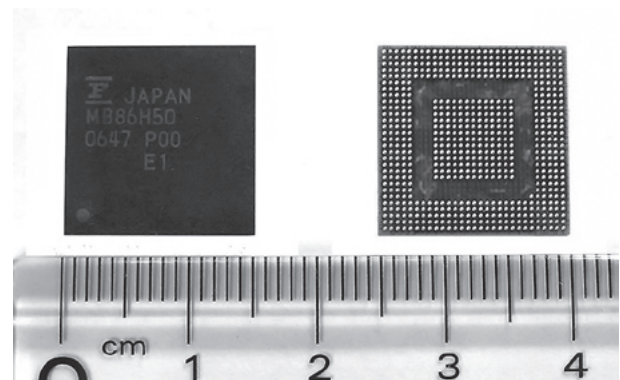


Figure 2  
MB86H50 chip photograph.

values are used to adjust the vertical position to be prefetched next. In this way, a high hit ratio can be maintained even for global motion in the vertical direction such as a vertical pan.

External-memory data transfer rates with and without prefetch memory control are shown in **Figure 3**. In addition to data transfer rates for the P frame (indicated by symbols ● and ▲), the figure also shows data transfer rates for the B frame to which bi-prediction mode can be applied (indicated by symbols ◆ and ■). Because the B frame normally requires twice as many reference images as the P frame, the data-transfer-rate reduction effect of the former is lower than that of the latter for the same prefetch memory capacity. Nevertheless, with reductions of 40 to 45% for the P frame and 25 to 30% for the B frame, as determined by actual measurements, a significant reduction effect is

obtained in both cases. As a result, total power consumption reduction of about 10% is achieved for the entire LSI including system, audio, and FCRAM processing. We have confirmed that the LSI operates at an industry-leading low power value of 600 mW in typical encoding (1440 × 1080 i).

### 5. Reference block grouping

This section explains the reference block grouping technique developed to reduce access to external memory during the decoding process. To achieve more precise motion estimation, the H.264/AVC standard allows for the allocation of motion vectors to block sizes smaller than that of MPEG-2 and for the definition of high-precision motion vectors in quarter-pel units. Here, the smallest block size to which a motion vector can be allocated is 4 × 4 pixels. To generate a predict-

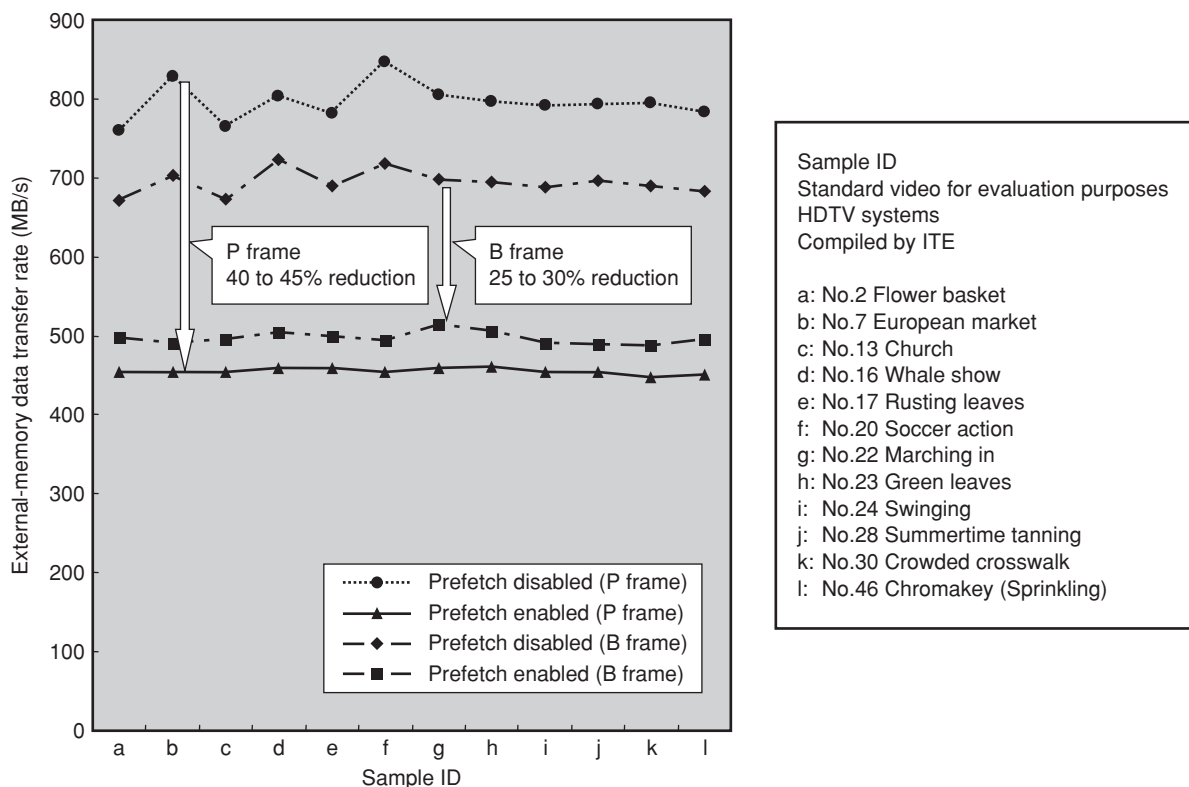


Figure 3 Data rates between LSI and external memory with and without prefetch technique.

ed image at quarter-pel precision, H.264/AVC provides for the reading of  $9 \times 9$ -pixel reference images and 6-tap finite impulse response filtering. This corresponds to the reading of large blocks, each more than five times the area of a small  $4 \times 4$ -pixel block. Accordingly, the process of reading reference images in H.264/AVC consumes a significant amount of external-memory bandwidth compared with MPEG-2.

This problem is solved by using the reference-block-grouping technique shown in **Figure 4** in the basic codec block. For the sake of simplicity, the example shown in the figure depicts a  $16 \times 16$ -pixel macroblock divided into four small  $8 \times 8$ -pixel blocks, each of which has

been allocated a motion vector. In this technique, after computing the motion vector for each block, the system calculates a reference-block group that contains all of the four reference blocks pointed to by the motion vectors. Now, if the size of the reference-block group is less than a certain threshold value, the system reads in that reference-block group from external memory as a reference image. On the other hand, if the size of the reference-block group is greater than that threshold, the system determines that the reference-block group contains many unnecessary data and therefore decides to read in the four reference areas individually. In the example shown here, the amount of data corresponding to the reading in of individual reference images is

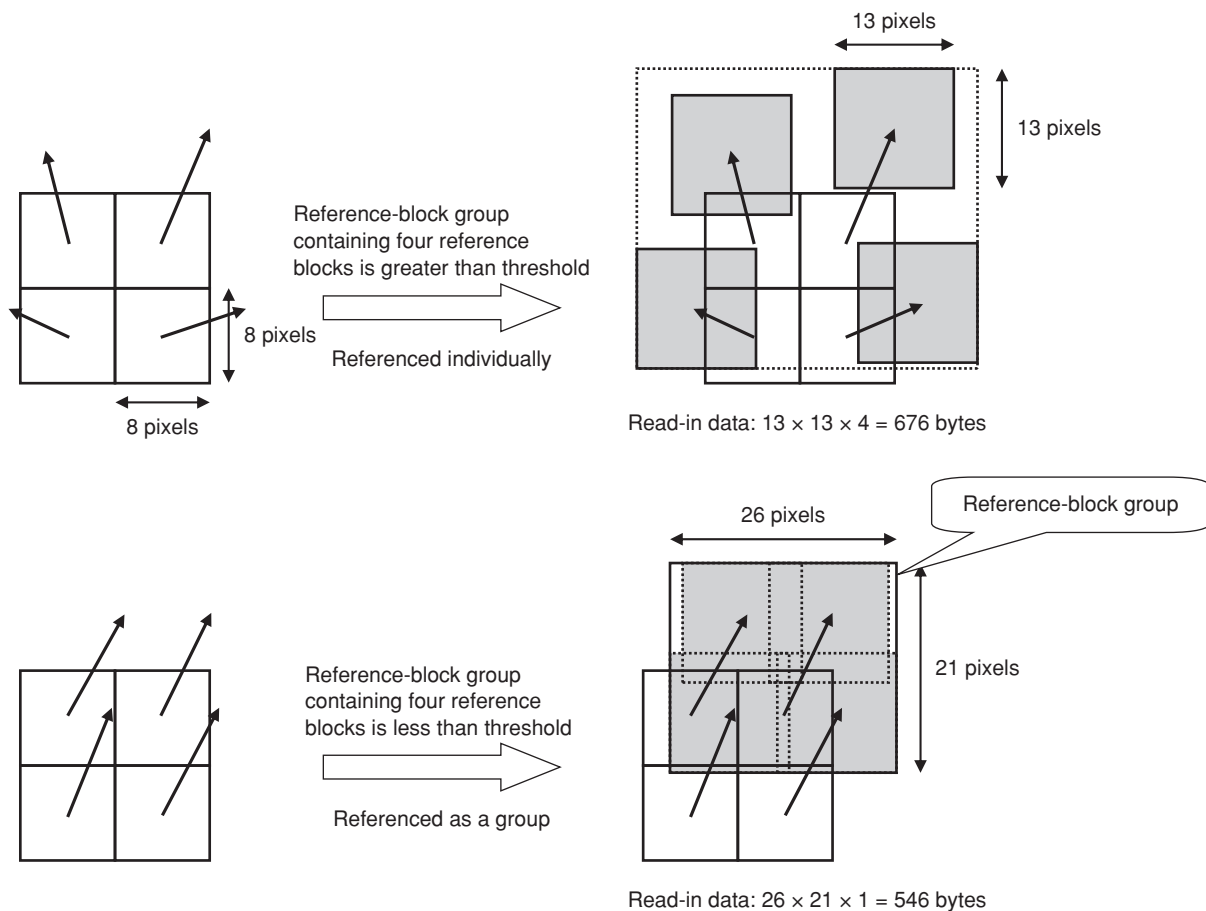


Figure 4 Example of reference-block-grouping technique.

676 bytes while that of a reference-block group is only 546 bytes. This means that reference-block grouping results in data reduction of about 20%. It can also reduce the number of times that external memory needs to be accessed. Overall, the technique raises the command efficiency of external-memory control in the memory controller and improves the effective bandwidth of external memory. **Figure 5** shows the data transfer rate (bar graphs) and number of reference read requests (indicated by ▲ and ■) with and without the grouping technique for the same video evaluation streams presented in Figure 3. It was found by actual measurements that the reference-block-grouping technique could reduce the number of reference read requests by 10 to 30% and the amount of data transferred from external memory by up to 7%.

## 6. Conclusion

This paper outlined a video codec LSI for HDTV video conforming to the H.264/AVC standard, which is slated to be the next-generation encoding standard. Particular attention was paid to the ability of this LSI to reduce external-memory access. The MB86H50 has already been incorporated into consumer AV devices and Fujitsu's IP-9500 IP video transmission equipment. Customers have given high praise for the superior picture quality and low power consumption of these products.<sup>4)</sup>

Combined with technical innovations such as high-speed wireless environments, advanced security techniques, digital rights management, and high-density storage media, we foresee the coming of an era in which HDTV can be enjoyed anytime and anywhere by anyone. The MB86H50

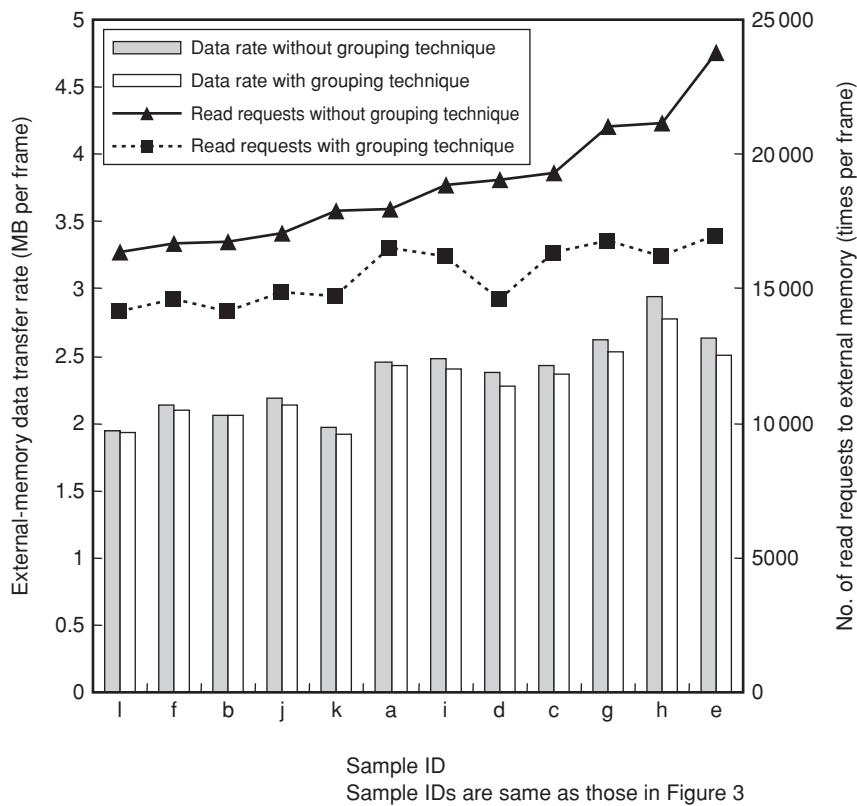


Figure 5  
Number of reference read requests and data rates with and without grouping technique.

LSI is at the vanguard of this trend. Fujitsu will continue to research and develop video-related LSI technology under the keywords of high picture quality and low power consumption.

## References

- 1) ISO/IEC 14496-10: Information technology — Coding of audio-visual objects — Part 10: Advanced Video Coding. 2005.



**Hiroshi Nakayama**

*Fujitsu Laboratories Ltd.*

Mr. Nakayama received the B.S. and M.S. degrees in Electronics Engineering from the University of Tokyo, Tokyo, Japan in 1985 and 1987, respectively. He joined Fujitsu Laboratories Ltd., Kawasaki, Japan in 1987 and has been engaged in research and development of 3D graphics systems and video codec

LSIs. He is a member of the Information Processing Society of Japan (IP SJ).

E-mail: h.nakayama@jp.fujitsu.com



**Yasuhiro Watanabe**

*Fujitsu Laboratories Ltd.*

Mr. Watanabe received the B.S. and M.S. degrees in Electronic Science and Engineering from Kyoto University, Kyoto, Japan in 1994 and 1996, respectively. He joined Fujitsu Laboratories Ltd., Kawasaki, Japan in 1996 and has been engaged in research and development of video codec LSIs.

E-mail: yas.watanabe@jp.fujitsu.com

- 2) A. Nakagawa: Fujitsu's Approach to H.264/AVC and Application Trends. *FUJITSU Sci. Tech. J.*, **44**, 3, p.343-350 (2008).
- 3) K. Kumagai et al.: System-in-Silicon Architecture and its Application to H.264/AVC Motion Estimation for 1080HDTV. *ISSCC Digest of Technical Papers*, p.1706-1715 (2006).
- 4) Y. Umezaki et al.: H.264/AVC Codec System for IP Networks. *FUJITSU Sci. Tech. J.*, **44**, 3, p.359-366 (2008).



**Akihiro Higashi**

*Fujitsu Laboratories Ltd.*

Mr. Higashi received the B.S. degree in Electronics and Communications Engineering from Meiji University, Tokyo, Japan in 1983. He joined Fujitsu Ltd., Kawasaki, Japan in 1983, where he was engaged in research and development of home information terminals and their ICs. He was transferred to Fujitsu Laboratories

Ltd., Kawasaki, Japan in 1993, where he has been engaged in research and development of ICs for image processing systems. From 1996, he researched design methodologies for system-on-a-chip technology. He is currently developing video codec LSIs. He is a member of the Institute of Image Information and Television Engineers (ITE).

E-mail: higashi@jp.fujitsu.com