Optimizing Video Compression for Intel® Digital Security Surveillance applications with SIMD and Hyper-Threading Technology

by Chew Yean Yam
Intel Corporation

Abstract

The growing consciousness of the general public, government, and private industry drives the demand for tighter security. With the advent of new technologies, Digital Security Surveillance (DSS) is becoming an essential part of daily life. Today's sophisticated surveillance solutions require scalability to support high numbers of video channels, higher throughput at the input/output (I/O) interface, processor performance to handle sophisticated algorithms, and reliable high-volume data storage.

Video compression is one of the most computationally demanding modules in DSS applications. Optimization in software can reduce the video compression burden on the processor, freeing processor resources for other important applications, including video analysis and graphical user interfaces. This article provides an introduction to the video compression performance improvements that can be achieved by features enabled in current Intel® Architecture Processors.

Introduction

The current generation of IA-32 processors, including Intel® Pentium® 4 processor, Intel® Xeon™ processor, and Intel® Pentium® M processor, includes multiple features designed to enhance performance. These include Single Instruction Multiple Data (SIMD) instruction set extensions, Hyper-Threading Technology (HT Technology),[1] microarchitectures that enable instruction execution at high clock rates, and a high-speed cache hierarchy.

Single Instruction Multiple Data (SIMD)

The SIMD computational technique was introduced in the IA-32 architecture with MMX™ technology and then further enhanced with Intel's introduction of Streaming SIMD Extensions (SSE), Streaming SIMD Extensions 2 (SSE2), and Streaming SIMD Extensions 3 (SSE3). SSE, SSE2, and SSE3 extend MMX technology by providing additional processor instructions for single-precision and double-precision floating-point operations, in addition to adding 128-bit registers and instructions to improve synchronization between multi-threaded agents, Table 1.

SIMD extensions improve the performance of applications characterized by i) *inherent parallelism*, ii) *recurring memory access patterns*, iii) *localized recurring operations performed on the data*, and iv) *data-independent control flow*. SIMD addresses the needs of multimedia, communications, and graphics/video applications used in DSS. These applications often use sophisticated algorithms to perform the same operations repetitively on large amounts of data. Such applications include 3-D graphics rendering, 3-D geometry, image processing, and video encoding and decoding.

Computing without SIMD is known as scalar processing, in which one instruction is performed on one data point at a time. Computation with SIMD enables the processor to execute one instruction on multiple data points concurrently, which increases the amount of data that can be processed in a given time interval, Figure 1.

MMX™ Technology	SSE	SSE2	SSE3
64-bit MMX™ Technology Registers	128-bit XMM Registers	128-bit XMM Registers	128-bit XMM Registers
Packed byte, word, double word integers	Packed single-precision floating-point operands	Packed double-precision floating-point operands	Floating-point instructions for asymmetric and horizontal computation and for thread synchronization.

Table 1—The next generations of Simple Instruction Multiple Data (SIMD), Streaming SIMD Extensions (SSE), SSE2, and SSE3 are extensions of Intel's original MMX technology. They provide additional instructions, enabling single-precision and double-precision floating-point operations, in addition to adding 128-bit registers and instructions to improve synchronization between multi-threaded agents.

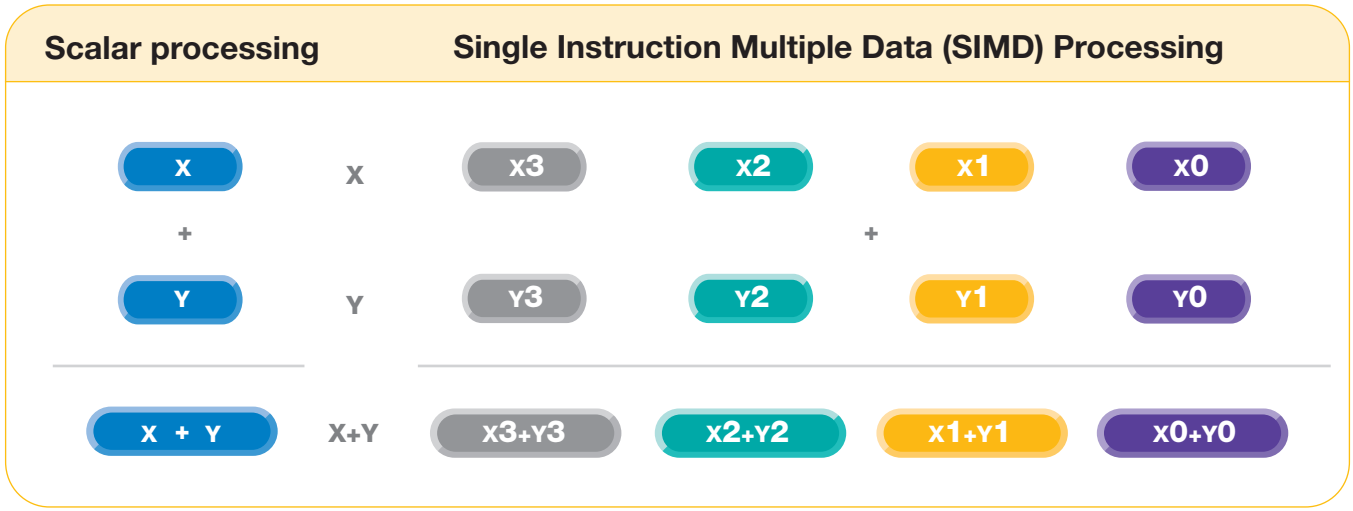


Figure 1—Comparison between computing with and without Single Instruction Multiple Data (SIMD) technique. SIMD improves performance over scalar processing by enabling the concurrent execution of one instruction on multiple data points.

Hyper-Threading Technology (HT Technology)

HT Technology enables software to take advantage of task-level or thread-level parallelism by providing multiple logical processor resources within a physical processor package. The operating system and applications perceive the single physical processor as two logical processors by maintaining the architectural state of these two processors. HT Technology provides more efficient multitasking and system responsiveness by allocating processor resources to applications as needed, enabling the processor to complete more tasks at a given time, and by efficiently sharing resources that would otherwise be idle, Figure 2. HT Technology can provide a significant performance gain in many current DSS applications, such as those that involve multiple video data streaming. This technology is supported by specific members of Intel Pentium 4 processor and Intel Xeon processor product lines.

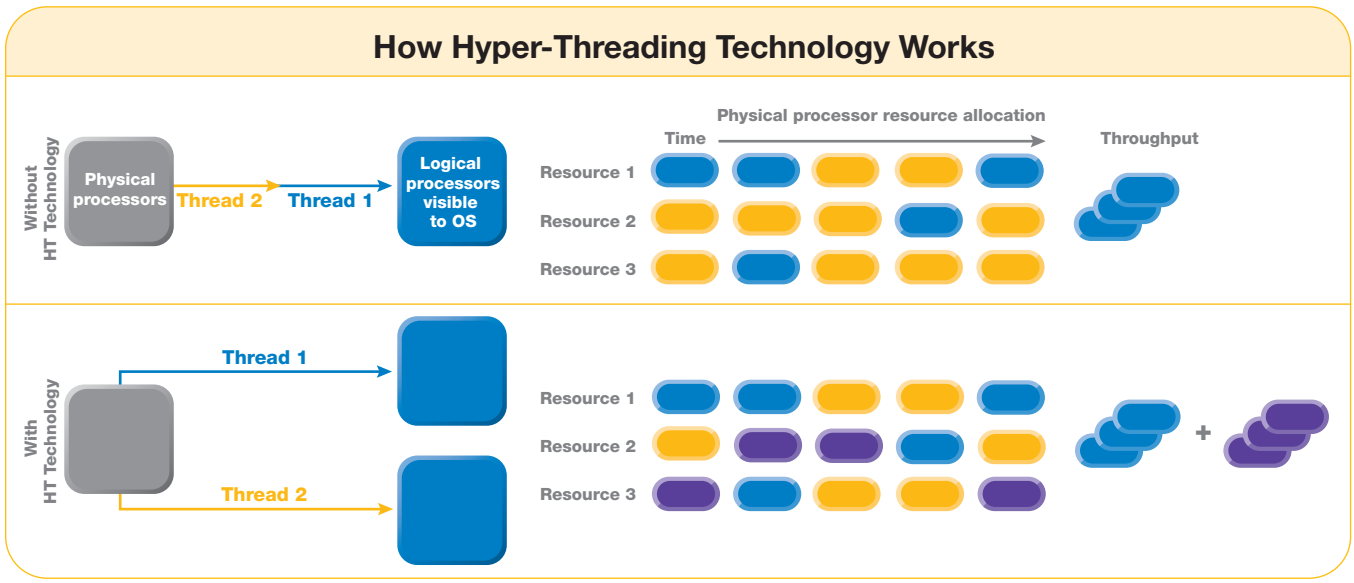


Figure 2—Hyper-Threading Technology boosts performance by sharing otherwise idle resources.

Video Encoding

Figure 3 illustrates the video encoding process, which is an essential function in many DSS applications. The highlighted modules are computationally intensive functions that can be optimized with the SIMD technique using MMX technology, SSE, and SSE2 technology.

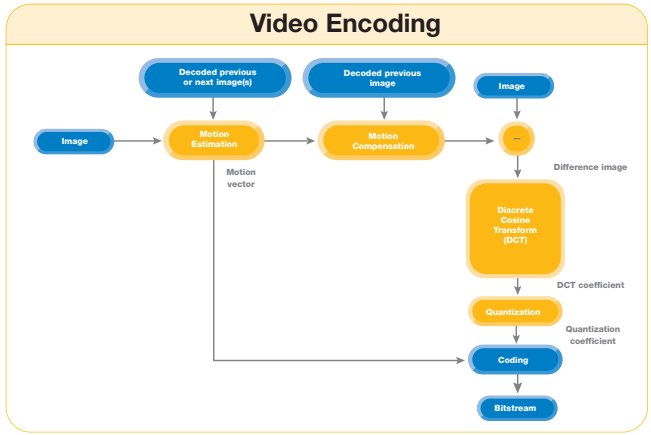
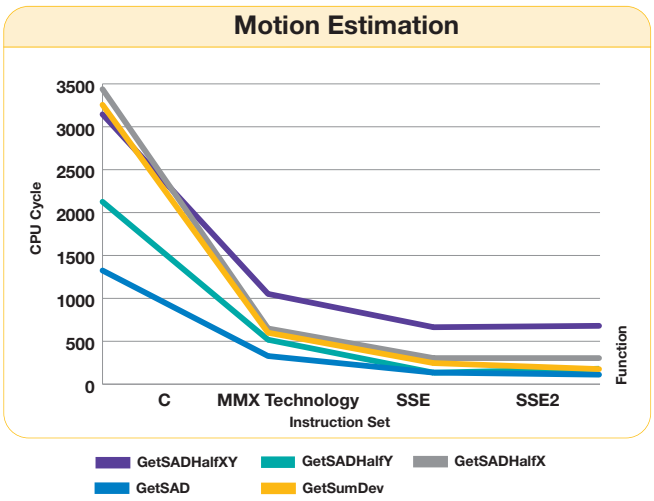
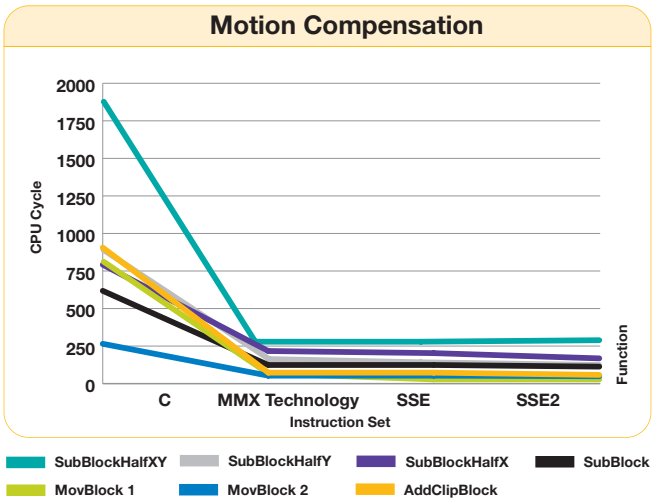
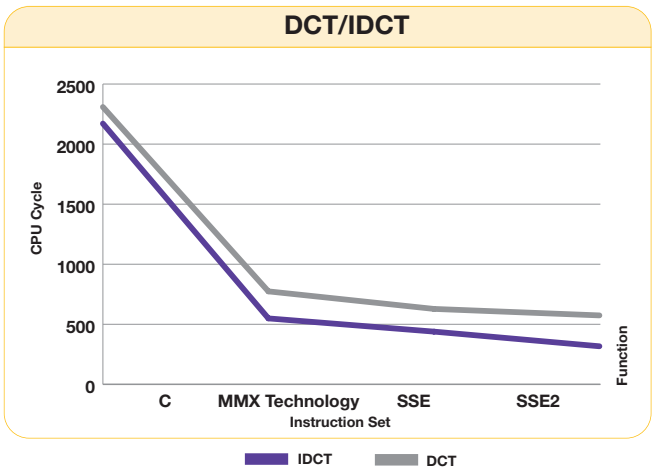


Figure 3—The video encoding process, with highlighted areas that show where SIMD-based computation can accelerate application performance.

Measuring Performance Gains







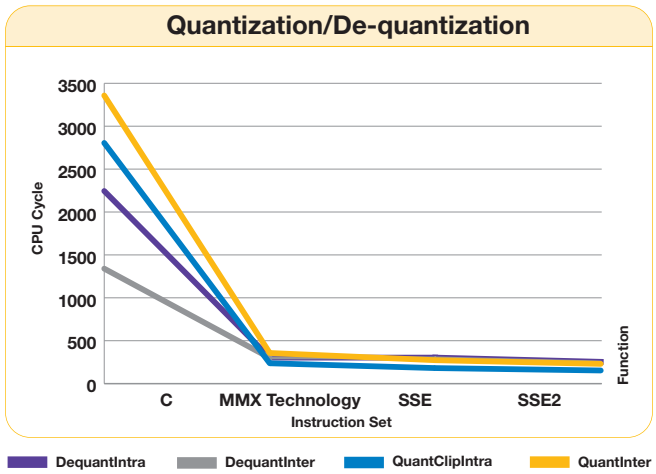


Figure 4—Graphs illustrate processor cycles required for motion estimation, motion compensation, Discrete Cosine Transform (DTC)/ Inverse Discrete Cosine Transform (IDCT), and quantization/de-quantization. The colored bars provide a comparison of processor cycles required when the functions are implemented in C language and when they are optimized using Intel® MMX™ technology, SSE, and SSE2. Source: Huper Laboratory Co., Ltd.

All the tests are run on an Intel Pentium 4 processor 2.4 GHz with Intel® 915G Express chipset-based platform. The first experiment observed the number of clock cycles required by executing the functions implemented in C, MMX technology, SSE, and SSE2 within the 4 major modules in video compression. The test is conducted by executing the functions for 10,000,000 times and the average is obtained to ensure a precise value. As shown in Figure 4, the number of processor cycles needed to perform video encoding functions drops dramatically with the use of later generations of SIMD instruction sets. More operations can be carried out in the same amount of time, resulting in higher performance. The significant drop in processor cycles correlates with a significant performance boost compared to C-based computation. The SUM function provides a general indication of the performance gain for each module. Results from motion estimation, motion compensation, DCT/IDCT, and quantization/de-quantization all exhibit this general trend.

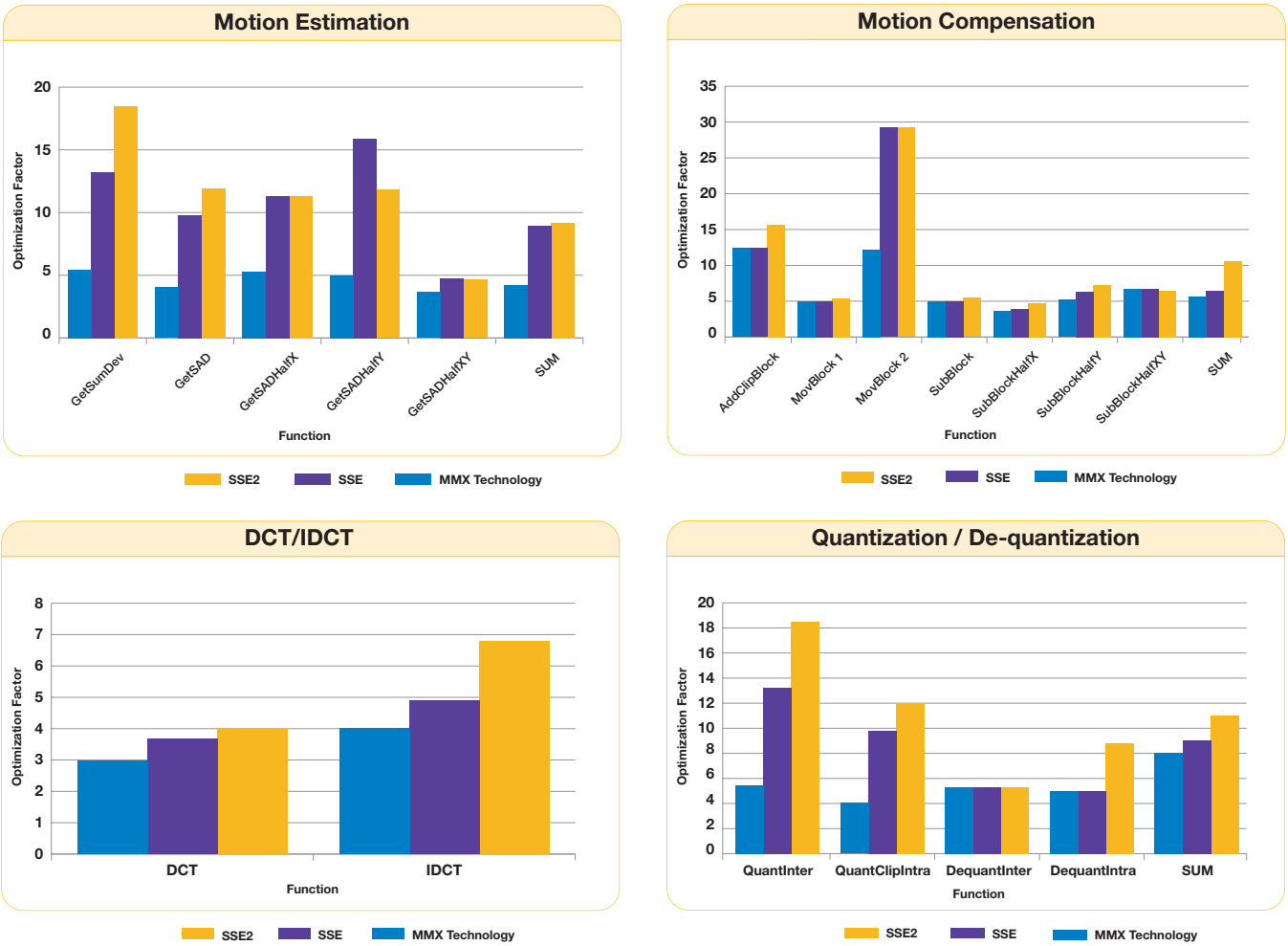
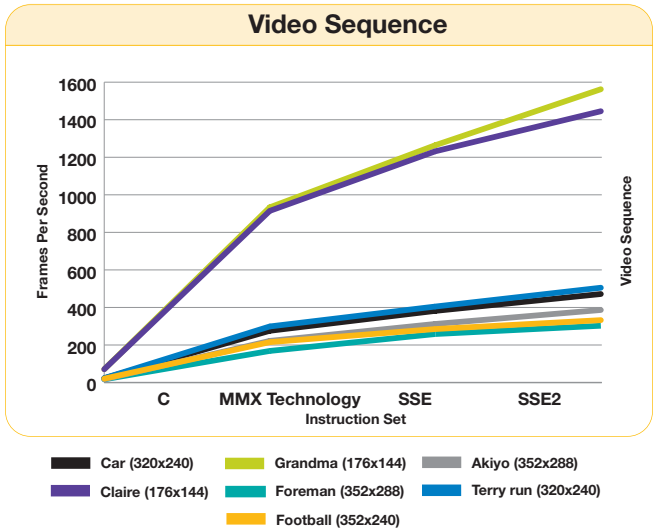


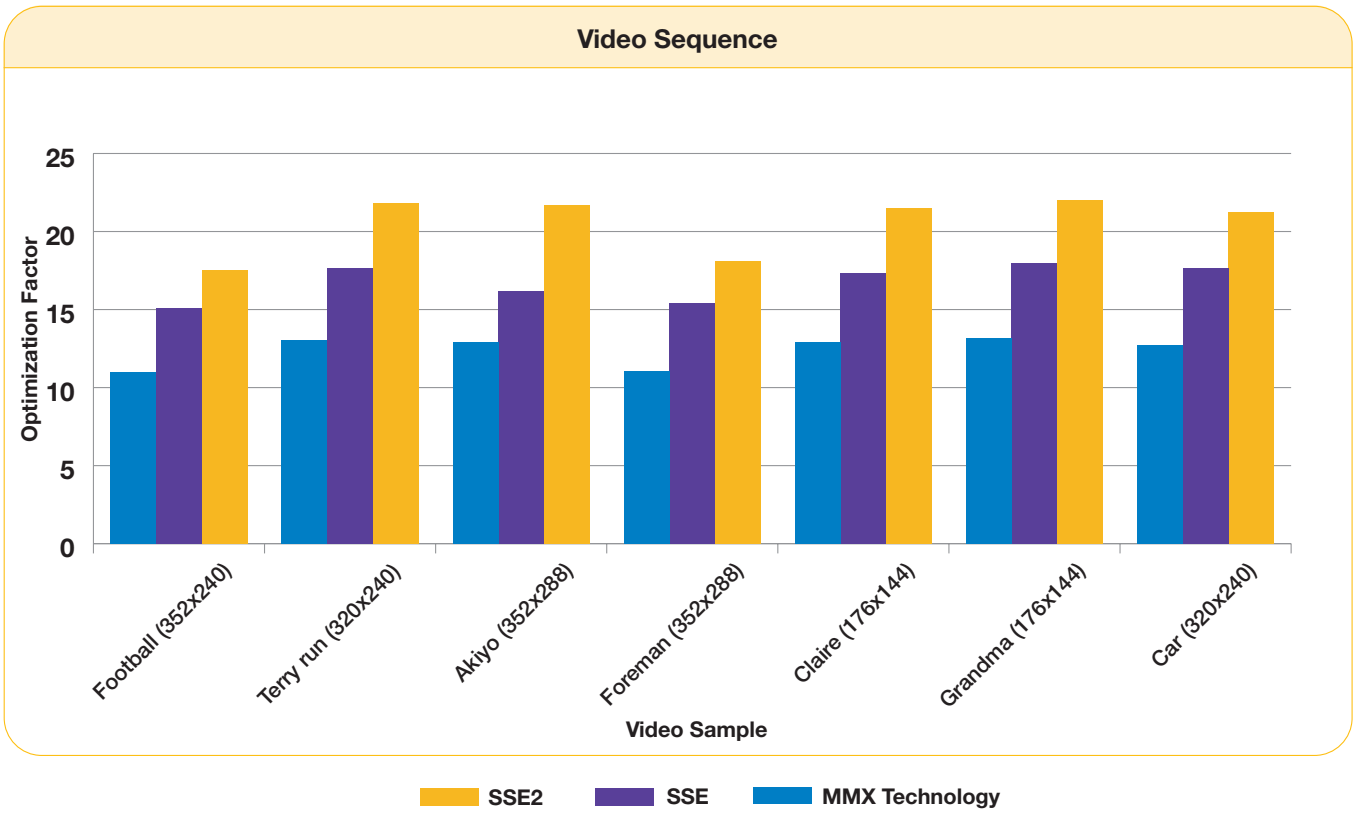
Figure 5—These graphs illustrate the optimization factor for motion estimation, motion compensation, DCT/IDCT, and quantization/de-quantization when optimized with MMX technology, SSE, and SSE2, compared to implementation in C language.

As shown in Figure 5, the optimization factor is the ratio of processor cycles needed to carry out the same task in a C-based implementation compared to an implementation with MMX technology, SSE, and SSE2. The performance improvement is proportional to the optimization factor. For example, the function GetSAD in motion estimation, when optimized by MMX technology, shows a 5x improvement compared to the C language implementation. SSE provides a 10x improvement, and the improvement for SSE2 is 12x. The SUM function provides a general indication of the performance gain for a particular module.

In the second section, graphs shown describe the number of frames per second the test program can handle. Figure 6 shows optimization techniques applied to real-life video samples. Total frames per second increase with the application of later generations of SIMD instruction set extensions, providing a significant performance improvement compared to a C-based implementation. MMX technology, SSE, and SSE2 provide an optimization factor of approximately 12, 17, and 20 respectively.



a. Frames per second achieved by C and SIMD instruction sets implementation on video encoding.



b. Improvement achieved by various samples of video clips when implemented in C and SIMD technique.

Figure 6—Effects of SIMD technique when applied to real-life video (resolutions indicated in parentheses).

This section illustrates the improvement achieved by implementing the test program with or without Hyper-Threading Technology feature. The application of HT Technology to a real video sample significantly increased the frame rate. The actual number of frames processed depends largely on scene complexity and the amount of motion involved in the scene. These examples show a performance improvement of 7–20 percent (Table 2).

Video Sequence (resolution)	HT Technology Disabled	HT Technology Enabled	= Enabled/Disabled
Football (352x240)	332.44	381.95	1.15
Terry run (320x240)	505.05	545.12	1.08
Akiyo (352x288)	386.53	415.78	1.08
Foreman (352x288)	302.59	333.11	1.10
Claire (176x144)	1,445.19	1,740.31	1.20
Grandma (176x144)	1,562.40	1,785.22	1.14
Car (320x240)	471.56	506.56	1.07

Table 2—Effects of HT Technology on a real video sample, expressed in frames per second.

Conclusion

These results show that the features available in the current generation of IA-32 processors enhance the processing efficiency of the major functional modules involved in video encoding. The application of MMX technology, SSE, and SSE2 instruction set extensions results in significant performance improvements compared to C-based implementations. The greatest gains were achieved by SSE2. When applied to real-life video samples, the application of MMX technology and later SIMD computational enhancements resulted in respective performance increments of 12, 17, and 20 times, while HT Technology resulted in a 7–20 percent improvement.

For developers of DSS solutions, these performance gains translate to improved processor utilization and the benefit of significantly enhanced application performance.

[1]Hyper-Threading Technology requires a computer system with an Intel® Pentium® 4 processor supporting Hyper-Threading Technology and an HT Technology–enabled chipset, BIOS, and operating system. Performance will vary depending on the specific hardware and software you use. See www.intel.com/info/hyperthreading/ for more information, including details on which processors support HT Technology.

Performance tests and ratings are measured using specific computer systems and/or components and reflect the approximate performance of Intel products as measured by those tests. Any difference in system hardware or software design or configuration may affect actual performance. Buyers should consult other sources of information to evaluate the performance of systems or components they are considering purchasing. For more information on performance tests and on the performance of Intel products, visit http://www.intel.com/performance/resources/limits.htm.

0905/SAC/MRM/PP/1K 309629-001