Image Processing Examples Using the TMS320C64x Image/Video Processing Library (IMGLIB)

Chris Chung
Oliver Sohm

TMS320C6000 Software Applications

ABSTRACT

The TMS320C64x™ image/video processing library (IMGLIB) provides a set of C-callable, assembly-optimized functions commonly used in imaging applications. While IMGLIB can be used to develop practical real-time applications with its high performance and ease of use, there are several important factors to consider in a system environment for attaining optimal performance mainly due to the data-intensive nature of imaging applications. This application report presents the usage and performance of several IMGLIB functions to help users utilize IMGLIB in their system development, and also presents performance analysis with three kinds of memory scenarios to help understand potential overhead related to memory hierarchy.

Contents

1 Introduction ................................................................. 2
2 Benchmarking ............................................................. 3
  2.1 Emulation/Simulation Setup ....................................... 3
  2.2 Cycle Count Measurement ......................................... 5
  2.3 Example Scenarios and Expected Performance .................. 6
    2.3.1 Scenario 1: Data in L1D ..................................... 6
    2.3.2 Scenario 2: Data in L2 SRAM ................................. 7
    2.3.3 Scenario 3: Data in Off-Chip Memory ........................ 7
  2.4 Data Alignment .................................................. 10
3 Examples ................................................................. 10
  3.1 Histogram .......................................................... 10
  3.2 Threshold .......................................................... 11
  3.3 Dithering .......................................................... 12
  3.4 Correlation ....................................................... 13
4 References ............................................................. 13

List of Figures

Figure 1. Memory Hierarchy and Potential Overhead .................. 2
Figure 2. Linker Command File for Scenarios 1 and 2 ................ 6
Figure 3. Linker Command File for Scenario 3 ........................ 8
Figure 4. Simplified Double Buffering Code ........................... 9

TMS320C64x is a trademark of Texas Instruments.
Trademarks are the property of their respective owners.
1 Introduction

Imaging applications are often considered as data-intensive, as well as compute-intensive, because they typically process large amount of data in real time, requiring efficient data transfer mechanisms as well as high computing power. TMS320C64x is an advanced, very long instruction word (VLIW) processor, suitable for imaging applications with its high computing power and large on-chip memory. It also provides enhanced direct memory access (EDMA) and cache to efficiently transfer data to/from off-chip memory. To help TMS320C64x users shorten the time-to-market in system development, Texas Instruments provides a set of assembly-optimized key imaging functions, named imaging/video processing library (IMGLIB). Each function in the IMGLIB is designed to produce the best performance possible by optimally utilizing available resources and avoiding potential resource conflicts. Therefore, when developing a system utilizing IMGLIB, it is important to understand potential overhead related to memory hierarchy, in order to estimate and improve the actual performance of a system being developed.

Figure 1 shows the memory hierarchy of C64x™ and related potential overhead. For example, without considering compulsory misses, when the program is bigger than the size of the level-one program cache (L1P), L1P cache misses can occur, stalling the central processing unit (CPU) until the required code is fetched. Similarly, when the data do not fit in the level-one data cache (L1D), L1D cache misses stall the CPU. All L1P and L1D misses are serviced by the level-two cache/static random-access memory (L2 cache/SRAM).

Figure 1. Memory Hierarchy and Potential Overhead

C64x is a trademark of Texas Instruments.
Similar to the L1D misses, L2 cache misses occur if the code and data do not fit in the L2 cache. The L2 miss overhead can be significant, compared to the L1P/L1D miss overhead, because the L2 cache needs to communicate with slow off-chip memory via EDMA. Considering the growing speed disparity between the processor and off-chip memory, careful data transfer handling (e.g., reducing L2 misses) is one of critical factors for attaining higher performance.

L2 SRAM can also be used to service L1D/L1P misses. However, EDMA is required to transfer code/data between L2 SRAM and off-chip memory if the code and data do not fit in the L2 SRAM. The data transfer with EDMA is typically more effective than that with L2 cache due to its nature of longer burst transactions, reducing memory access latency overhead. However, the EDMA transfer involves more programming effort because data transfers and synchronization have to be manually managed. TMS320C64x provides both cache and EDMA mechanisms to allow users to choose the right mechanism, depending on situations.

This application report presents the usage and performance of several IMGLIB functions to help TMS320C64x users utilize IMGLIB in system development. In addition, performance analysis with three kinds of memory scenarios is presented to help understand potential overhead related to memory hierarchy.

2 Benchmarking

2.1 Emulation/Simulation Setup

A TMS320C6416 test and evaluation board (TEB) is used in this application report to measure cycle counts for IMGLIB examples. Table 1 lists key features of the C6416 TEB, which are important factors in performance analysis and optimization. More details on the C6416 internal memory structure and operations can be found in *TMS320C64x DSP Two-Level Internal Memory Reference Guide* (SPRU610).

<table>
<thead>
<tr>
<th>Item</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>C6416</td>
<td>Clock frequency 500 MHz</td>
</tr>
<tr>
<td>L1P</td>
<td>16K-byte, direct-mapped, 32-byte cache line</td>
</tr>
<tr>
<td>L1D</td>
<td>16K-byte, 2-way set associative, 64-byte cache line</td>
</tr>
<tr>
<td>L2 SRAM</td>
<td>8-cycle L1P miss penalty, 6-cycle L1D miss penalty, up to 1M byte</td>
</tr>
<tr>
<td>L2 cache</td>
<td>8-cycle L1P/D miss penalty, up to 256K bytes, 4-way set associative, 128-byte cache line</td>
</tr>
<tr>
<td>L2 to L1D read path</td>
<td>256 bits</td>
</tr>
<tr>
<td>L1D to L2 write buffer</td>
<td>64-bit, merge with 4 outstanding write misses</td>
</tr>
<tr>
<td>EMIF</td>
<td>EMIF-A: 64-bit bus, EMIF-B: 32-bit bus</td>
</tr>
<tr>
<td>SDRAM</td>
<td>Clock frequency 133 MHz</td>
</tr>
<tr>
<td>Bus width</td>
<td>64 bits (two 32-bit modules) connected to EMIF-A</td>
</tr>
<tr>
<td>Page size</td>
<td>2K bytes</td>
</tr>
</tbody>
</table>
The C6416 TEB is connected to a PC through an XDS560 board, and a Code Composer Studio™ Integrated Development Environment (IDE) configuration, based on “_C64xx XDS560 Emulator Address 0”, is used as follows. If you use other types of interfaces, e.g., XDS510, please make sure to choose the right configuration.

Code Composer Studio is a trademark of Texas Instruments.
Be sure to select the right General Extension Language (GEL) file for the C6416 TEB.

If you use simulation, select “C6416 Fast Sim Ltl Endian”. The cycle counts obtained from simulation might not be accurate, especially with off–chip memory accesses.

Software version numbers used in this application report are as follows:

- Code Composer Studio: version 2.1
- C64 IMGLIB: version 1.02b

**2.2 Cycle Count Measurement**

The built-in timer in C6416 is used to measure cycle counts for IMGLIB examples. The following sample code shows how to set up the timer and measure cycle counts with Chip Support Library (CSL).

```c
hTimer = TIMER_open(TIMER_DEVANY, 0);  /* open a timer */

/* Configure the timer. 1 count corresponds to 8 CPU cycles in C64 */
TIMER_configArgs(hTimer, 0x000002C0, 0xFFFFFFFF, 0x00000000);

/* Compute the overhead of calling the timer. */
start = TIMER_getCount(hTimer); /* to remove L1P miss overhead */

/* Call a function here. */
diff = (TIMER_getCount(hTimer) - start) - overhead;

printf("%d cycles \n", diff*8);
```

The maximum resolution of the timer is 8 CPU cycles, since the input clock to the timer is fixed to the CPU clock divided by eight. The function call overhead for TIMER_getCount() is roughly measured and compensated. Additional information on the timer registers can be found in the *TMS320C6000 Peripherals Reference Guide* (SPRU190).
2.3 Example Scenarios and Expected Performance

Three kinds of scenarios are presented in this report, considering potential overhead related to memory hierarchy:

1. When data are in L1D
2. When data are in L2 SRAM
3. When data are in off-chip memory.

In these examples, L1P miss overhead can be ignored because the examples are small enough to fit L1P. However, L1P misses have to be carefully managed in the case where L1P thrashing overhead is significant.

2.3.1 Scenario 1: Data in L1D

This scenario is shown to validate the formula cycle counts listed in the TMS320C64x Image/Video Processing Library Programmer's Reference (SPRU023). Note that the formula cycle count assumes flat memory, not considering any overhead related to memory hierarchy. Figure 2 shows a linker command file used for this scenario. For more information on linker commands, refer to the TMS320C6000 Optimizing Compiler User's Guide (SPRU187).

Information on TMS320C6000 memory maps can be found in the TMS320C6000 Assembly Language Tools User's Guide (SPRU186).

Since all code and data are initially stored in L2SRAM, the example function is called twice, to eliminate L1P/L1D miss overhead, and a cycle count for the second call is measured.

```
MEMORY
{
   L2SRAM:   o = 00000000h   l = 00100000h  /* 1 Mbytes */
}
SECTIONS
{
   .cinit      >       L2SRAM
   .text       >       L2SRAM
   .stack      >       L2SRAM
   .bss        >       L2SRAM
   .const      >       L2SRAM
   .data       >       L2SRAM
   .far        >       L2SRAM
   .switch     >       L2SRAM
   .sysmem     >       L2SRAM
   .tables     >       L2SRAM
   .cio        >       L2SRAM
}
```

Figure 2. Linker Command File for Scenarios 1 and 2

TMS320C6000 is a trademark of Texas Instruments.
2.3.2 Scenario 2: Data in L2 SRAM

In this scenario, L1D miss overhead needs to be considered. The linker command file for Scenario 1 (shown in Figure 2) is used for this scenario. Table 2 lists expected stall cycles related to L1D read and/or write transactions. When there are read transactions only, the number of stall cycles (without considering pipelined misses) is the number of L1D read misses times L1D miss penalty (i.e., 6 cycles). In case of write transactions only, there is no stall unless the write buffer is full.

When there are both read and write transactions, the L1D read miss penalty can increase because any write transaction in the write buffer has to be flushed before a read miss is serviced to maintain data coherency.

Table 2. Stall Cycles Related to L1D

<table>
<thead>
<tr>
<th>Transaction</th>
<th>Number of Stall Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Read transaction only</td>
<td>Number of L1D read misses * L1D miss penalty</td>
</tr>
<tr>
<td>Write transaction only</td>
<td>No stall cycle unless the write buffer is full</td>
</tr>
<tr>
<td>Read and write transactions</td>
<td>Number of L1D read misses * (L1D miss penalty + additional cycles for write buffer flush)</td>
</tr>
</tbody>
</table>

2.3.3 Scenario 3: Data in Off-Chip Memory

When data are in off-chip memory, either (1) L2 cache or (2) L2 SRAM with EDMA is used to transfer the data in off-chip memory. When L2 cache is used, there will be L2 cache miss overhead as well as L1D miss overhead. In addition, the L1D read miss penalty is higher with L2 cache (i.e., 8 cycles) than with L2SRAM (i.e., 6 cycles). On the other hand, the off-chip memory access has to be manually managed with L2 SRAM, whereas it is seamlessly handled by the cache controller with L2 cache.

EDMA is typically advantageous over cache in terms of performance for the following three reasons. First, the overhead of access latency with EDMA is less than that with cache, since a EDMA transfer can be much longer than a cache line transfer. Second, computation and data transfers can be tightly overlapped with EDMA, which often results in significant performance improvement. Third, the L2 write-allocate policy can result in more data transfers than needed (i.e., load/allocate/writeback instead of load/store).

With EDMA, the overall processing time depends on the ratio between the compute time and the data transfer time, as categorized into two conditions listed in Table 3. In the compute-bound condition where the compute time is greater than the data transfer time, the overall processing time is determined by the compute time. Note that the compute time is defined as the processing time without the overhead of off-chip memory accesses. The data transfer time is defined as the time to transfer data to/from off-chip memory, which consists of memory access latency and burst transfer time.

Table 3. Compute-Bound vs. Memory-Bound

<table>
<thead>
<tr>
<th>Condition</th>
<th>Description</th>
<th>Overall Processing Time</th>
</tr>
</thead>
<tbody>
<tr>
<td>Compute-bound</td>
<td>Compute time &gt; data transfer time</td>
<td>Compute time + time for the first load and last store transfers</td>
</tr>
<tr>
<td>Memory-bound</td>
<td>Compute time &lt; data transfer time</td>
<td>Data transfer time + EDMA management overhead</td>
</tr>
</tbody>
</table>
To reduce the access latency overhead in memory accesses, the burst transfer must be long enough. However, too long burst transfers can cause negative effect in compute-bound cases because time for the first load and last store transfers cannot be hidden behind the compute-time. In the memory-bound condition where the compute time is less than the data transfer time, the overall processing time is determined by the data transfer time; thus further improving the compute time does not contribute to higher performance.

From an experiment on the C6416TEB, the cycle count of 34,000 was measured in transferring 64K bytes of data between on–chip and off–chip memories. This cycle count will be used to analyze EDMA examples in Section 3.

Figure 3 shows a linker command file used for this scenario.

```c
MEMORY
{
  L2SRAM:  o = 00010000h  l = 000F0000h  /* 960 kbytes */
  CE0:     o = 80000000h  l = 01000000h  /* 16 Mbytes */
}
SECTIONS
{
  .cinit        >       L2SRAM
  .text         >       L2SRAM
  .stack        >       L2SRAM
  .bss          >       L2SRAM
  .const        >       L2SRAM
  .data         >       L2SRAM
  .far          >       L2SRAM
  .switch       >       L2SRAM
  .sysmem       >       L2SRAM
  .tables       >       L2SRAM
  .cio          >       L2SRAM
  .imgbuf       >       CE0  /* User created data section for off–chip memory */
}
```

**Figure 3. Linker Command File for Scenario 3**

The following statements are used to allocate image arrays to a user-defined section (.imgbuf) in off-chip memory.

```c
#pragma DATA_SECTION(in_image, ".imgbuf")
#pragma DATA_SECTION(out_image, ".imgbuf")
```

Figure 4 shows a double buffering code used for EDMA transfers, which utilizes the DAT module in CSL. The double buffering code is simplified for easier explanation, thus it can only handle the case where the total data size is a multiple of the buffer size times two. It uses two sets of input and output buffers, called InBuffA/OutBuffA and InBuffB/OutBuffB. With the two sets of buffers, the CPU can process data with one set of buffers while EDMA transfers data in another set of buffers.
The DAT_wait() is used to wait for a transfer to complete. The DAT_copy() issues a data transfer that happens in the background of CPU processing. The first two blocks of input data are transferred to input buffers before the double buffering loop begins. Once the transfers have completed, you can process the data in InBuffA and store the results to OutBuffA. The data in OutBuffA is sent to off-chip memory, and the next DAT_copy() begins copying the next input block to InBuffA for future processing. While these two transfers are occurring, you process the data in InBuffB and store the results to OutBuffB. When it is done, the result in OutBuffB is sent to off-chip memory.

```c
number_of_transfers = total_data_size / buffer_size;
/*------------------- Initial transfers -------------------------------*/
id_InBuffA  = DAT_copy(in_image, InBuffA, buffer_size);
id_InBuffB  = DAT_copy(in_image + buffer_size, InBuffB, buffer_size);
/* Begin Double Buffering */
for(i=0; i < number_of_transfers; i+=2)
{
    DAT_wait(id_InBuffA);   /* wait for transfers to complete */
    DAT_wait(id_OutBuffA);
    /* ------------- Process the data in InBuffA and store the results to OutBuffA */
    Process( InBuffA, OutBuffA, buffer_size );
    id_OutBuffA = DAT_copy(OutBuffA, out_image + (i* buffer_size), buffer_size);
    if( i < number_of_transfers-2 )
        id_InBuffA  = DAT_copy(in_image + ((i+2)* buffer_size), InBuffA,
                                buffer_size);
    DAT_wait(id_InBuffA);
    DAT_wait(id_OutBuffB);
    
    /* --------------- Process the data in InBuffB and store the results to OutBuffB */
    Process( InBuffB, OutBuffB, buffer_size );
    id_OutBuffB = DAT_copy(OutBuffB, out_image + (i+1)* buffer_size],
                   buffer_size);
    if( i < number_of_transfers-2 )
        id_InBuffB  = DAT_copy(in_image + ((i+3)* buffer_size), InBuffB,
                                buffer_size);
}
```

**Figure 4. Simplified Double Buffering Code**
2.4 Data Alignment

Due to the structure of internal memory/cache, some IMGLIB functions require input/output memory arrays to be aligned to a specific boundary. This restriction must be carefully managed in all C64x devices for attaining optimal performance. As an example, the following statement is used to allocate the array (input) to a 8-byte boundary.

```c
#pragma DATA_ALIGN (input, 8)
```

The C64x compiler automatically aligns arrays of all types to an 8-byte boundary if they are not declared in a `struct` statement. When dynamic memory allocation is used, the allocated memory is also aligned to an 8-byte boundary, regardless of types. More information on data alignment rules by the compiler can be found in the TMS320C6000 Optimizing Compiler User’s Guide (SPRU187).

The structure of internal memory/cache on the C64x generation varies from device to device. Therefore, refer to the appropriate device data sheet to determine the structure of a particular device.

3 Examples

This section presents the usage and performance of few IMGLIB functions: histogram, threshold, dithering, and correlation. To minimize the variation in cycle count measurement, be sure to select the Reset menu (under Debug in Code Composer Studio) before running an example. The cycle counts listed in this section were measured in Release mode.

3.1 Histogram

Image histogram is used to count the number of occurrences of pixel intensity in an image, which is widely used in image analysis and enhancement. The `IMG_histogram` function, which computes the histogram on an 8-bit image, is defined as:

```c
void IMG_histogram (unsigned char * restrict in_data, int n, short accumulate, short * restrict t_hist, short * restrict hist)
```

The maximum number of pixels that can be counted in each bin is 65535. Note that a temporary array, `t_hist`, must have 1024 16-bit entries and be initialized to zero. For example, with EDMA, this function is called for a block of data. Therefore, `t_hist`, must be cleared before each call. The input array (in_data) must be aligned to a 4-byte boundary and the number of input data (`n`) must be a multiple of 8. Table 4 lists IMG_histogram benchmarks.

<table>
<thead>
<tr>
<th>Number of Data (N)</th>
<th>Formula (9/8) * N + 228</th>
<th>Data In On-Chip Memory</th>
<th>Data In Off-Chip Memory (500-MHz CPU, 133-MHz SDRAM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Scenario 1 (L1D)</td>
<td>Scenario 2 (L2 SRAM)</td>
</tr>
<tr>
<td>8,192</td>
<td>9,444</td>
<td>9,448</td>
<td>10,224†</td>
</tr>
<tr>
<td>65,536</td>
<td>73,956</td>
<td>–</td>
<td>80,344†</td>
</tr>
</tbody>
</table>

† Without further L1D optimization
‡ With 8K-byte EDMA buffers
The cycle count for Scenario 1 is close to the formula cycle count because no cache miss occurred. For Scenario 2, there will be L1D miss overhead (i.e., 6 cycles) for every 64 bytes of input data. For example, the expected number of cycles with 64K bytes of data is \(80,344 = 73,956 \text{ (i.e., formula cycles)} + 65,536 \text{ (i.e., number of data in bytes)} / 64 \text{ (i.e., L1D cache line)} \times 6 \text{ (i.e., L1D miss overhead with L2 SRAM)}\), which is close to the measured cycle count of 80,344.

Since the cycle count for Scenario 2 (i.e., 80,344) is larger than the transfer cycles (i.e., 34,000 from section 2.3.3) in this example, this function is considered as compute-bound with EDMA; thus, expected performance includes time for the first and last block transfers as well as the performance of Scenario 2. Another overhead to consider is the time to clear \(t_{hist}\) eight times instead of one, which increases the overall processing time since this function is compute-bound.

### 3.2 Threshold

Image threshold has many different uses in image/video processing systems, including converting grayscale images to binary images for morphological processing, and converting image formats suitable for segmentation. The IMGLIB contains four threshold functions:

- **IMG_thr_gt2max**: pixels greater than the threshold are set to 255.
- **IMG_thr_gt2thr**: pixels greater than the threshold are set to the threshold.
- **IMG_thr_le2min**: pixels less than or equal to the threshold are set to 0.
- **IMG_thr_le2thr**: pixels less than or equal to the threshold are set to the threshold.

The four functions are defined as:

```c
void IMG_thr_gt2max (const unsigned char * restrict in_data, unsigned char * restrict out_data, short cols, short rows, unsigned char threshold)
void IMG_thr_gt2thr (const unsigned char * restrict in_data, unsigned char * restrict out_data, short cols, short rows, unsigned char threshold)
void IMG_thr_le2min (const unsigned char * restrict in_data, unsigned char * restrict out_data, short cols, short rows, unsigned char threshold)
void IMG_thr_le2thr (const unsigned char * restrict in_data, unsigned char * restrict out_data, short cols, short rows, unsigned char threshold)
```

Note that the number of data is specified in two-dimension variables, \(cols\) and \(rows\). The input (\(in\_data\)) and output (\(out\_data\)) arrays must be aligned to an 8-byte boundary and must not be overlapped. The number of input data (\(cols * rows\)) must be at least 16 and a multiple of 16.

Table 5 lists IMG_gt2max benchmarks.

### Table 5. IMG_thr_gt2max Benchmarks

<table>
<thead>
<tr>
<th>Number of Data (N)</th>
<th>Formula (0.1875 * N + 22)</th>
<th>Data In On-Chip Memory</th>
<th>Data In Off-Chip Memory (500-MHz CPU, 133-MHz SDRAM)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Scenario 1 (L1D)</td>
<td>Scenario 2 (L2 SRAM)</td>
</tr>
<tr>
<td>8,192</td>
<td>1,558</td>
<td>1,568</td>
<td>2,856†</td>
</tr>
<tr>
<td>65,536</td>
<td>12,310</td>
<td>–</td>
<td>22,576†</td>
</tr>
</tbody>
</table>

† Without further cache optimization
‡ With 16K-byte EDMA buffers
The cycle count for Scenario 1 is close to the formula cycle count because no cache miss occurred. For Scenario 2, there will be L1D write miss overhead as well as L1D read miss overhead for every 64 bytes of data. For example, the expected number of cycles with 64K bytes of data is 18,454 cycles \([18,454 = 12,310 \text{ (i.e., formula cycles)} + 65,536 \text{ (i.e., number of data in bytes)} / 64 \text{ (i.e., L1D cache line)} \times 6 \text{ (i.e., L1D read miss penalty with L2 SRAM)}]\), which shows about 18% error compared to the measured cycle count of 22,576. This is because L1D read miss stall cycle count increases when the write buffer is not empty.

Since the cycle count of Scenario 2 (i.e., 22,576) is much less than the transfer cycles (i.e., 68,000) in this example, this function is considered as memory-bound with EDMA; thus, expected cycle count includes the data transfer cycles and EDMA management overhead. In this case, further improving the compute time does not contribute to higher performance.

### 3.3 Dithering

Image dithering is commonly used in printing applications. The C64 IMGLIB dithering function (IMG_errdif_bin) implements the Floyd-Steinberg error diffusion algorithm, which is defined as:

```c
void IMG_errdif_bin (unsigned char * restrict errdif_data, int cols, int rows, short * restrict err_buf, unsigned char thresh)
```

The `err_buf[ ]` array must be initialized to 0 prior to the first call. Each pixel in the `errdif_data[ ]` is compared against a user-specified threshold (thresh). Pixels larger than the threshold are set to 255, while other pixels are set to 0. The error value for a pixel is propagated to the neighboring pixels using the Floyd-Steinberg filter. The number of columns (`cols`) must be at least 2. Table 6 lists IMG_errdif_bin benchmarks.

**Table 6. IMG_errdif_bin Benchmarks**

<table>
<thead>
<tr>
<th>Number of Data (N = \text{COLS} \times \text{ROWS})</th>
<th>Formula ((4 \times \text{COLS} + 11) \times \text{ROWS} + 7)</th>
<th>Data In On-Chip Memory</th>
<th>Data In Off-Chip Memory (500-MHz CPU, 133-MHz SDRAM)</th>
<th>Number of Cycles</th>
</tr>
</thead>
<tbody>
<tr>
<td>8,192 = 128 \times 64</td>
<td>33,479</td>
<td>33,488</td>
<td>Scenario 1 (\text{L1D})</td>
<td>271,184†</td>
</tr>
<tr>
<td>65,536 = 256 \times 256</td>
<td>264,967</td>
<td>–</td>
<td>Scenario 2 (\text{L2 SRAM})</td>
<td>286,616‡</td>
</tr>
</tbody>
</table>

† Without further cache optimization
‡ With 4K-byte EDMA buffers

The cycle count for Scenario 1 is close to the formula cycle count due to no cache misses. For Scenario 2, there will be L1D miss overhead (i.e., 6 cycles) for every 64 bytes of input data. For example, the expected number of cycles with 64K bytes of data is 271,111 cycles \([271,111 = 264,967 \text{ (i.e., formula cycles)} + 65,536 \text{ (i.e., number of data in bytes)} / 64 \text{ (i.e., L1D cache line)} \times 6 \text{ (i.e., L1D miss overhead with L2 SRAM)}]\), which is close to the measured cycle count of 271,184. There is no additional stall cycle related to write buffer full, since the input memory is used as the output memory (i.e., in-place computation).

Since the cycle count for Scenario 2 (i.e., 271,184) is larger than the transfer cycles (i.e., 68,000) in this example, this function is considered as compute-bound with EDMA; thus, expected cycle count includes cycles for the first and last block transfers as well as the cycles of Scenario 2.
3.4 Correlation

Image correlation is used for image matching. The IMGLIB correlation function (IMG_corr_3x3) computes correlation of an image with a 3x3 mask. The output will have the highest value at the best-matched input image location. The IMG_corr_3x3 is defined as:

```c
void IMG_corr_3x3 (const unsigned char * restrict in_data, int * restrict out_data, const unsigned char mask[3][3], int x_dim, int n_out)
```

This function needs to be called once for each row. However, it may be invoked for multiple rows at a time by setting the number of output (n_out) to a multiple of the width (x_dim). In this case, two outputs at the end of each row will have meaningless values, thus care must be taken when interpreting the results. All arrays (in_data, out_data and mask) must not be overlapped. The number of outputs (n_out) must be a multiple of 8. Table 7 lists IMG_corr_3x3 benchmarks.

<table>
<thead>
<tr>
<th>Number of Cycles</th>
<th>Data In On-Chip Memory</th>
<th>Data In Off-Chip Memory</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Scenario 1 (L1D)</td>
<td>Scenario 2 (L2 SRAM)</td>
<td>Scenario 3 (EDMA)</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Transfer Cycles</td>
</tr>
<tr>
<td>8,192</td>
<td>12,310</td>
<td>13,232††</td>
</tr>
<tr>
<td></td>
<td>12,320</td>
<td>—</td>
</tr>
<tr>
<td>65,536</td>
<td>98,326</td>
<td>105,560††</td>
</tr>
<tr>
<td></td>
<td></td>
<td>184,504††</td>
</tr>
</tbody>
</table>

† Without further cache optimization
‡ With 4K-byte input and 16-kbyte output EDMA buffers

In Scenarios 1 and 2, this example is similar to the image threshold in section 3.2, since it has separate input and output arrays.

The IMG_corr_3x3 function requires N+2 rows of input data to compute N rows of output results. Due to the nature of block processing with EDMA, input image blocks are overlapped by 2 rows. Therefore, the size of input data becomes 73,728 bytes [73,728 = 65,536 + (65,536/4,096)*2*256] and the total size of data transfers is 335,872 bytes (335,872 = 73,728 + 65,536*4), which corresponds to 174,250 cycles (174,250 = 34,000 * 335,872 / 65,536). Since the cycle count of Scenario 2 (i.e., 105,560) is less than the transfer cycles (i.e., 174,250) in this example, this function is considered as memory-bound; thus, expected cycle count includes the data transfer cycles and EDMA management overhead.

4 References

1. TMS320C64x DSP Two-Level Internal Memory Reference Guide (SPRU610).
2. TMS320C6000 Peripherals Reference Guide (SPRU190).
3. TMS320C64x Image/Video Processing Library Programmer’s Reference (SPRU023).
IMPORTANT NOTICE

Texas Instruments Incorporated and its subsidiaries (TI) reserve the right to make corrections, modifications, enhancements, improvements, and other changes to its products and services at any time and to discontinue any product or service without notice. Customers should obtain the latest relevant information before placing orders and should verify that such information is current and complete. All products are sold subject to TI's terms and conditions of sale supplied at the time of order acknowledgment.

TI warrants performance of its hardware products to the specifications applicable at the time of sale in accordance with TI's standard warranty. Testing and other quality control techniques are used to the extent TI deems necessary to support this warranty. Except where mandated by government requirements, testing of all parameters of each product is not necessarily performed.

TI assumes no liability for applications assistance or customer product design. Customers are responsible for their products and applications using TI components. To minimize the risks associated with customer products and applications, customers should provide adequate design and operating safeguards.

TI does not warrant or represent that any license, either express or implied, is granted under any TI patent right, copyright, mask work right, or other TI intellectual property right relating to any combination, machine, or process in which TI products or services are used. Information published by TI regarding third–party products or services does not constitute a license from TI to use such products or services or a warranty or endorsement thereof. Use of such information may require a license from a third party under the patents or other intellectual property of the third party, or a license from TI under the patents or other intellectual property of TI.

Reproduction of information in TI data books or data sheets is permissible only if reproduction is without alteration and is accompanied by all associated warranties, conditions, limitations, and notices. Reproduction of this information with alteration is an unfair and deceptive business practice. TI is not responsible or liable for such altered documentation.

Resale of TI products or services with statements different from or beyond the parameters stated by TI for that product or service voids all express and any implied warranties for the associated TI product or service and is an unfair and deceptive business practice. TI is not responsible or liable for any such statements.

Mailing Address:
Texas Instruments
Post Office Box 655303
Dallas, Texas 75265

Copyright © 2003, Texas Instruments Incorporated