

WIRELESS MPEG-4 VIDEO ON TEXAS INSTRUMENTS DSP CHIPS

Madhukar Budagavi, Wendi Rabiner, Jennifer Webb, Raj Talluri

DSP Solutions R&D Center, Texas Instruments, Incorporated, MS 8374
8330 LBJ Freeway, Dallas, TX 75243

madhukar@ti.com, wendi@mit.edu, webb@ti.com, talluri@ti.com

ABSTRACT

Technology has advanced in recent years to the point where multimedia communicators are beginning to emerge. These communicators are low-power, portable devices that can transmit and receive multimedia data through the wireless network. Due to the high computational complexity involved and the low-power constraint in wireless applications, these devices require the use of processors that are powerful and are at the same time very power-efficient. In order to facilitate interoperability, it is important that these devices use standardized compression and communication algorithms. As a first step in implementing multimedia terminals, Texas Instruments (TI) has demonstrated real-time MPEG-4 video decoding (simple profile) on a TMS320C54x, TI's low power, high performance DSP chip. In addition, TI has outlined a system-level solution to transmitting video across wireless networks, including channel coding and communication protocols.

1. BACKGROUND

As wireless telephony is becoming commonplace, more and more features are being supported on wireless networks. By the year 2001 [1], wireless multimedia communicators will likely become a reality. Standards such as MPEG-4 [2] are being finalized just as the required processing power is becoming affordable [3]. A key component in multimedia communicators is the video codec which enables video communication. In MPEG-4, the video coding part of the standard encompasses a range of applications, including a wireless ("simple") profile.

The small size of wireless communicators limits the display size available for video, yet even a small display is adequate for many applications. For instance, an SQCIF format (128×96 pixels) can be useful for videoconferencing, surveillance, news, or entertainment. In addition, a smaller format requires less bandwidth, less memory (18 kbytes per frame for 4:2:0 SQCIF), and coding artifacts may be less objectionable. Higher resolution will be possible as bandwidth increases and display technology improves.

By combining MPEG-4 simple video compression with the error resilient H.223 communication algorithms [4], it is possible to design a wireless video compression system that achieves good quality reconstructed video. Such a system is depicted in Fig. 1. Low-cost real-time video decoding has only recently become feasible. On the 40 MHz TMS320C541, it is possible to decode thirty SQCIF frames per second for relatively simple sequences

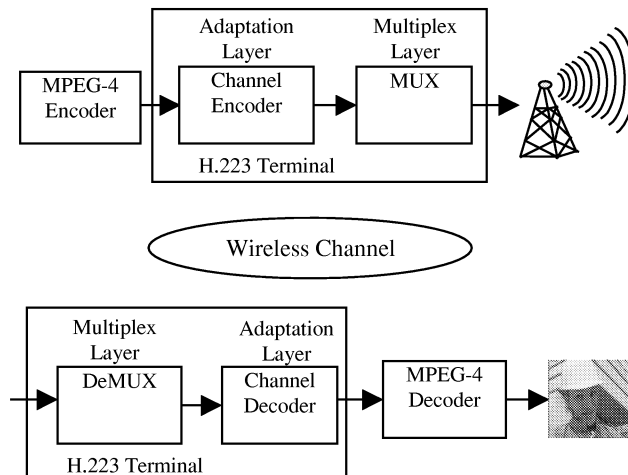


Figure 1. Wireless video transmission system

with little motion, e.g., talking heads, coded around 20 kbps. In the near future, DSP chips will be capable of providing the processing power necessary for more complex multimedia applications.

In the next section, we give an overview of the TMS320C54x DSP family and highlight the salient features that have made it ideally suited for wireless applications. In Section 3, we expand on the wireless video communication system outlined in Fig. 1 and describe the H.223 multiplex and the MPEG-4 video simple profile, which includes error-resilience tools, in more detail.

2. TMS320C54x DSP CHIPS [5]

Currently, about half of all cell phones have a TI DSP chip inside. The TMS320C54x family of DSP processors from Texas Instruments gives designers an edge in today's wireless and wireline communications markets. This generation provides the right combination of ultra-low power dissipation, high performance, and cost effectiveness to meet the needs of a variety of terminal and infrastructure communications equipment.

The 'C54x DSP chips are designed to execute up to 100 million instructions per second (MIPS) while dissipating as little as 0.45 mW/MIPS. This power-efficient performance is demanded by

wireless applications such as cellular phones, pagers, PCS terminals, and portable information appliances. With up to 100 MIPS performance, the highly efficient 'C54x architecture helps to enable emerging wireless multimedia applications.

2.1. Innovative Architecture

The high performance levels of the TMS320C54x DSP chips are made possible by an innovative architecture designed to meet the needs of a variety of applications (see Fig. 2). This architecture reduces Viterbi "butterfly update" operations down to only four instruction cycles for GSM channel decoding. This frees MIPS so the device CPU can perform other system tasks. Other key features include:

- four internal buses and dual address generators enable multiple operand operations to reduce memory bottlenecks,
- a 40-bit adder and two 40-bit accumulators support parallel instructions that execute in only one instruction cycle,
- a second 40-bit adder available at the output of the multiplier allows unpipelined MAC operation as well as dual addition and multiplication in parallel,
- single-cycle normalization and exponential encoding support floating-point arithmetic,
- new single-cycle instructions efficiently execute common DSP tasks like a symmetrical FIR filter,
- a 40-bit arithmetic logic unit (ALU) features a dual 16-bit configuration capability to enable dual, one-cycle operations, and
- eight auxiliary registers and a software stack enable the industry's most advanced fixed-point DSP C compiler.

2.2. Low Power Device

A key limitation to wireless communications is power consumption. When a DSP chip executes a task more efficiently, designers are able to use remaining MIPS to implement tasks normally handled by off-chip ASICs or microcontrollers. This integration not only yields space savings, but power savings as well. Also, when a DSP task can be done more efficiently, the DSP chip can spend more time in power-down, or IDLE, mode.

The TMS320C54x DSP chips feature three IDLE modes and low voltage operation down to 2.5 V at full performance. In the IDLE modes, the 'C54x DSP chip enters a dormant state and dissipates considerably less power than in normal operation. The efficiency of the TMS320C54x architecture enables TI to provide the lowest milliwatts per function capability for the wireless communications market.

2.3 High Processing Performance

To support the processing and memory requirements of a variety of wireless multimedia terminals, the C54x family of processors is available in a number of different configurations. The 'C541 features 5K words of RAM and 28K words of ROM. The 'C545 and 'C546 come with the addition of larger amounts of on-chip RAM, ROM, and intelligent peripherals. The 'C542 has 2K words of ROM and 10K words of RAM which provides the flexibility for many algorithms to be implemented on chip.

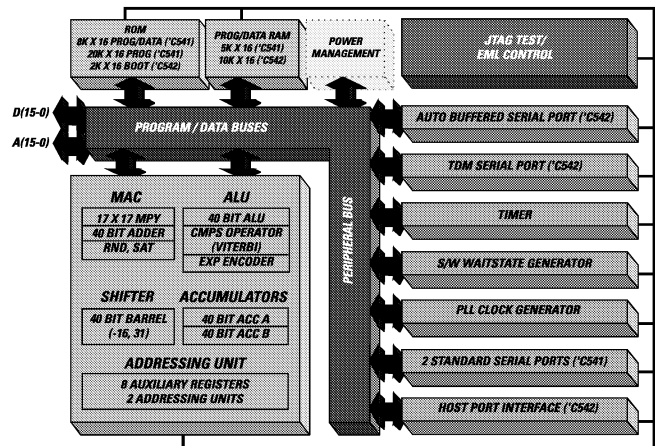


Figure 2. TMS320C54x architecture.

The newer 'C549 features 32K words of on-chip RAM and 16K words of ROM. To allow even higher integration levels and to further reduce chip count, power dissipation, and system cost, the 'C5410 provides twice the amount of on-chip RAM as that of the 'C549. In addition to the 64K words of RAM, the 'C5410 also has a six-channel DMA controller.

In addition to these devices, as the DSP chips move to the newer silicon process nodes, the C54x family of DSP chips provides a compelling road map of backward-compatible devices supplying hundreds of MIPS with very little power dissipation to address the needs of emerging wireless multimedia terminals.

3. WIRELESS VIDEO COMMUNICATION

Encoded video data are particularly sensitive to bitstream errors, due to the extensive use of variable-length coding (VLC) and temporal prediction. VLC causes the decoder to easily lose synchronization with the encoder in the presence of bit errors, and temporal prediction causes errors in the reconstructed video to propagate to future frames. Error correction techniques, such as forward error correction (FEC), can be used to reduce the number of errors in the bitstream at the cost of increased overhead.

As part of the H.223 multiplex standard [4], an adaptation layer may be used to provide additional protection from channel errors, beyond the level of service available from the network provider. In addition, the MPEG-4 video standard includes several error-resilience tools [6] for improved performance in the presence of channel errors.

3.1 Channel Coding in Adaptation Layer

To protect data against the harsh conditions typically present on wireless channels, the adaptation layer of the H.223 standard provides support for forward error correction (FEC) using Rate Compatible Punctured Convolutional (RCPC) encoding of the

data. The amount of protection can be set based on the channel conditions and the amount of allowed overhead to bring the aggregate bit error rate down to a level at which the MPEG-4 error resilience tools can be effective and provide acceptable quality at the decoder.

The FEC coded video data are sent to the multiplex layer. The multiplex layer performs multiplexing of the video, audio, and control data. In addition, the multiplex layer adds a resynchronization flag and a header to the multiplexed data (the payload). This flag is chosen so that it has good auto-correlation properties and has low cross-correlation with the data in the payload. Detection of the resynchronization flag is done at the H.223 decoder using correlation and thresholding. This allows a high degree of detection and a low degree of false detection in the presence of channel errors. The header added by the H.223 multiplex layer contains the length of the payload and a code into a multiplex table which tells the decoder how to demultiplex the video, audio, and data. This header is protected using an extended Golay error correction code. Fig. 3 shows the structure of an H.223 packet.

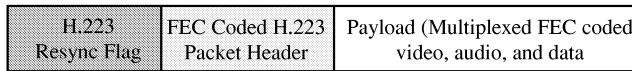


Figure 3. H.223 packet.

The H.223 packets are sent over a wireless channel, such as a GSM or DECT channel. These are bandwidth-constrained, error-prone channels. At the receiver, the (possibly corrupted) packets are demultiplexed and FEC decoded using the multiplex and adaptation layers of H.223, respectively. The FEC decoding is performed using a maximum a-posteriori Viterbi decoder. The 'C54x family of chips have an integrated "Viterbi accelerator", a dedicated instruction set which can efficiently perform this computationally intensive operation. The 'C54x is well-known for its high performance in wireless communication systems.

The error-corrected video bitstream is sent to the source decoder. Since the bitstream may contain some residual errors, it is important to use the MPEG-4 error resilience tools [6, 7] described next.

3.2 MPEG4 Error Resilience Features

Due to the FEC tradeoff between overhead and bit error reduction, it may not be possible to correct all bit errors present in the bitstream. Thus, the video decoder must be able to decode bitstreams that contain some errors. Researchers at Texas Instruments have been very active in the area of error-resilient video coding, and have contributed to the development of error-resilience tools for the MPEG-4 wireless (simple) profile.

The MPEG-4 video coding standard incorporates several error resilience tools into the standard to make the compressed bitstream more robust to channel errors. These techniques are aimed at improving error localization, enabling greater data recovery, and enhancing error concealment.

3.2.1 Resynchronization

A video decoder that is decoding a corrupted bitstream can lose synchronization with the encoder due to the use of variable length codes. MPEG-4 adopted a resynchronization strategy referred to as the video packet approach. A video packet consists of (i) a resynchronization marker (RS), (ii) a video packet header, and (iii) macroblock data, as shown in Fig. 4. The resynchronization marker is a unique 17-bit code that cannot be emulated by the variable length codes used in MPEG-4. Whenever an error is detected in the bitstream, the video decoder jumps to the next resynchronization marker to establish synchronization with the encoder. The video packet header contains information that helps in restarting the decoding process and consists of the absolute macroblock number of the first macroblock in the video packet and the initial quantization parameter used to quantize the DCT coefficients in the packet. A third field, labeled HEC, is also included in the video packet header. Its use is discussed in a later section. The macroblock data part of the video packet consists of the motion vectors, DCT coefficients, and mode information for the macroblocks contained in the video packet.

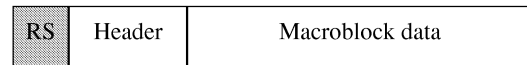


Figure 4. MPEG-4 video packet.

The encoder modifies the predictive encoding methods used such that there is no data dependency between adjacent video packets. This is required so that even if one of the video packets in the current image is corrupted due to errors, the others packets can be decoded and utilized by the decoder. The size of a video packet is not fixed by the MPEG-4 standard; however it is recommended that the size of the video packets (and hence the spacing between resynchronization markers) be approximately equal.

3.2.2 Data Partitioning

The data partitioning mode of MPEG-4 partitions the macroblock data within a video packet into a motion part and a texture part (DCT coefficients) separated by a unique Motion Marker (MM), as shown in Fig. 5. All the syntactic elements of the video packet that are required to decode motion related information are placed in the motion partition and all the remaining syntactic elements that relate to the DCT data are placed in the texture partition. If only the texture information is lost, data partitioning allows the use of motion information to conceal errors in a more effective manner.

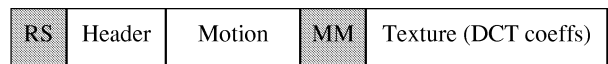


Figure 5. MPEG-4 data partitioned video packet.

The motion marker is computed from the motion VLC tables using a search program such that it is Hamming distance 1 from

any possible valid combination of the motion VLC table entries [8]. The motion marker indicates to the decoder the end of the motion information and the beginning of the DCT information. The number of macroblocks in the video packet is implicitly known after encountering the motion marker. Note that the motion marker is only computed once based on the VLC tables and is fixed in the standard. Based on the VLC tables in MPEG-4, the motion marker is a 17-bit word whose value is **1 1111 0000 0000 0001**.

3.2.3 Reversible Variable Length Codes (RVLCs)

Reversible VLCs are designed such that they can be instantaneously decoded both in the forward and the backward direction. While decoding the bitstream in the forward direction, if the decoder detects an error it jumps to the next resynchronization marker and starts decoding the bitstream in the backward direction until it encounters an error. Based on the two error locations, the decoder can recover some of the data that would have otherwise been discarded. This is illustrated in Fig. 6, which shows only the texture part of the video packet; only data in the shaded area is discarded. Note that if RVLCs were not used, all the data in the texture part of the video packet would have to be discarded. RVLCs thus enable the decoder to better isolate the error location in the bitstream.

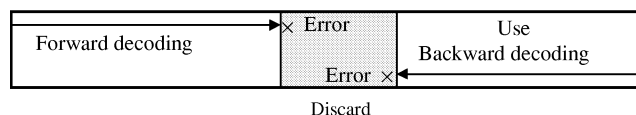


Figure 6. Use of RLVCs.

3.2.4 Header Extension Code (HEC)

Important information which remains constant over a video frame, such as the spatial dimensions of the video data, the time stamps associated with the decoding and the presentation of this video data, and the type of the current frame (INTER coded/INTRA coded), are transmitted in the header at the beginning of the video frame data. If some of this information is corrupted due to channel errors, the decoder has no other recourse but to discard all the information belonging to the current video frame. In order to reduce the sensitivity of this data, a 1-bit field called HEC was introduced in the video packet header. When HEC is set, the important header information that describes the video frame is repeated in the bits following the HEC. This duplicate information can be used to verify and correct the header information of the video frame. The use of HEC significantly reduces the number of discarded video frames and helps achieve a higher overall decoded video quality.

3.2.5 Adaptive Intra Refresh (AIR)

Temporal propagation of errors can be stopped by using non-predictive INTRA coding. The procedure of forcefully encoding some macroblocks in a frame in INTRA mode to flush out possible errors is called INTRA refreshing. INTRA refresh is very effective in stopping the propagation of errors, but it comes at the cost of a large overhead; coding a macroblock in INTRA

mode typically requires many more bits when compared to coding the macroblock in INTER mode. Hence the INTRA refresh technique has to be judiciously used.

For areas with low motion, simple error concealment by just copying the previous frame's macroblocks works quite effectively. For macroblocks with high motion, error concealment is difficult. Since the high motion areas are perceptually the most significant, any persistent error in the high motion area becomes very noticeable. The AIR technique of MPEG-4 makes use of the above facts and INTRA refreshes the motion areas more frequently, thereby allowing the possibly corrupted high motion areas to recover quickly from errors.

4. A TOTAL SOLUTION

While this paper describes the error-resilience tools, channel coding, and processor capabilities required for wireless video applications, TI's efforts go beyond demonstrating the ability to decode video in real-time on a low-power TMS320C54x processor. To support wireless multimedia communication, TI has numerous other projects in areas such as antenna, receiver, modulator, and power-supply design. In addition to DSP chips, TI also offers a wide range of complementary mixed-signal and analog products [9].

REFERENCES

- [1] L. Robinson, "Japan's New Mobile Broadcast Company: Multimedia for Cars, Trains, and Hand-Helds," *Advanced Imaging*, Jul. 1998, pp. 18-22.
- [2] MPEG-4 Video Group, "Overview of the MPEG-4 Standard," ISO/IEC JTC1/SC29/WG11 N2323, Dublin, Ireland, July 1998.
<http://drogo.cse.lt.stet.it/mpeg/standards/mpeg-4/mpeg-4.htm>
- [3] J. Eyre and J. Bier, "DSP processors hit the mainstream," *Computer*, Aug. 1998, pp. 51-59.
- [4] ITU-T Recommendation H.223, "Multiplexing protocol for low bitrate multimedia communication," Annex C, Sep 1997
- [5] Texas Instruments official TMS320C54x www site:
<http://www.ti.com/sc/docs/dsp/products/c5000/c54x/index.htm>
- [6] R. Talluri, "Error-resilient video coding in the ISO MPEG-4 standard," *IEEE Communication Magazine*, June 1998.
- [7] M. Budagavi and R. Talluri, "Wireless video communications," chapter in *Mobile Communications Handbook*, 2 ed., Ed. J. Gibson, CRC Press, 1998, to appear.
- [8] R. Talluri et al., "Error concealment by data partitioning," *Signal Processing: Image Commun.*, 1998, to appear.
- [9] Texas Instruments www sites:
wireless: <http://www.ti.com/sc/docs/wireless/home.htm>
RF: <http://www.ti.com/sc/docs/rf/index.htm>
Power: <http://www.ti.com/sc/docs/wireless/intro/power.htm>
MSP: <http://www.ti.com/sc/docs/msp/pran/default.htm>