HEWLETT PACKARD

Interactive Video from Desktops to Settops

Frederick Kitson, Vasudev Bhaskaran, Deven Kalra
Computer Systems Laboratory
HPL-95-58
June, 1995

video, compression, MPEG, settops, desktops, graphics

Video is the component of multimedia that provides the most visual realism while placing the most stress on a computer system. The capture, processing, transmission, digital storage and display of video requires a delicate balance between available MIPs, MB/sec and MB of dynamic and static memory. Multimedia applications of the future will combine interactive video and graphics in new and exciting forms. This paper will also address some of the issues and innovations in engendering *media enabled* computer systems from conventional desktops such as workstations with modern RISC processors to the next generation of consumer computers known as "settops". To provide a specific example, a MPEG1 decoder that is capable of real-time playback of video and audio on HP's RISC-based workstations will be described. The desktop community is seeking "TV-like" functions such as surround sound and broadcast quality video while the home consumer desires "computer-like" interactivity and connectivity.

Internal Accession Date Only
To be presented at and published in the proceedings of the *NEC Symposium in Multimedia*, Tokyo, Japan, June 7, 1995.

1 Introduction

Multimedia is often defined in terms of data types such as video, audio, images, graphics, numbers and text. Today tools exist to manipulate and access text, for example, via word processors or numbers via spreadsheets. With the capabilities of current Digital Signal Processors, audio has also become a supported data type. In the next few years, full-motion video will achieve this status. To this end, computer designers are commissioned to architect the fundamental capabilities to capture, manipulate, store and transmit video. Because of the performance requirements necessary to achieve such *dexterity*, compression algorithms are presently mandatory. Compression and decompression support have therefore become enabling technology for multimedia systems. Fortunately some standards have gained popular support such as MPEG for motion video so that VLSI manufacturers and application developers can create interoperable systems.

As with audio, the first systems to support video have been realized with specialized processors that are tuned for DSP or video in particular. As general purpose processors achieve higher MIP ratings and adapt to this new media type, video processing will come under the domain of workstation applications. This paper will present an example of both situations. First a summary of HP's Precision Architecture (PA) RISC processor will be presented to show how a high performance general processor can support an efficient system for handling compressed video and audio. MPEG1 playback was chosen as a key goal since it is a good match in terms of complexity and current applications such as CD-ROM support.

This work can be extended to other operations on video such as scaling and merging of video streams for teleconferencing. New applications such as medical imaging can be entertained with support for 2D and 3D data at video or interactive rates (10-30 frames/sec.).

Consumer video systems will be a catalyst for the development of an aggressive price/performance point with the primary objective of video decompression in real time and general purpose multimedia support as a secondary requirement. This low cost interactive video possibility is engendered by the confluence of video compression, digital processors, memory integration and communications processing with cable TV infrastructure. The *settop* will provide the interface from the communications interface or cable to the video monitor or television. This interactive processing device will demodulate, decode, decrypt and decompress digital video and audio streams as well as process analog video. It will be the customer interface for viewing and service selection for such applications as movies-on-demand, music-on-demand, games-on-demand and home shopping. HP is creating a *consumers computer* that represents a "Trojan Horse" into the home for information access. It will contain high volume, low-cost media processors and interfaces that can span from settops to desktops. We are meeting this challenge with innovative algorithms and architectures to meet the high performance requirements and low cost. The back channel for interactivity is key to enhanced applications and services for the next generation settop. Video and multimedia servers will support settops through a client-server relationship.

2 MPEG1 Decompression on HP Workstations

The decoding of a MPEG1 bitstream as performed on HP's workstations[1] includes (a) system level decoding to extract the timing information and demultiplexing of the compressed video and audio streams, and (b) video decoding[2] to decompress the MPEG1 video data. Audio decoding is done as well but this aspect will not be covered here. Figure 1 shows a block diagram of the MPEG1 video decoder.

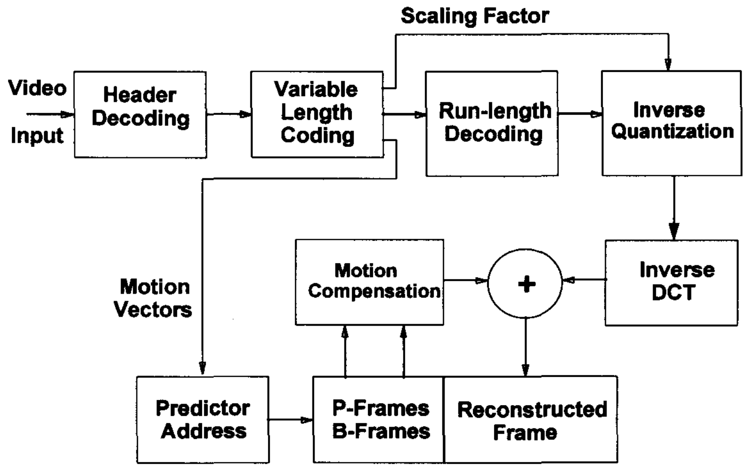

Figure 1 : MPEG1 Video Decoder

Video decoding consists of these steps:

1. Video sequence header decoding to extract parameters of the video sequence such as picture rate, bit rate, image size, etc. For each group of pictures (GOP), identify picture type, e.g. I, P or B picture. For each picture and for each slice within the picture, determine the quantizer scale.
2. For each slice, decode each macroblock. Macroblock layer decoding consists of extracting the motion-vectors from the coded stream and then extracting the DCT information for the blocks within the macroblock.
3. The DCT information is huffman coded. Thus huffman decoding is performed to decode the variable-length codes into fixed-length symbols.
4. Inverse quantization is performed on the huffman decoded data.
5. For each 8x8 block of the inverse quantized data, a 8x8 inverse DCT is computed. This transforms the data back to the image domain.
6. Motion-compensation is then performed if needed. For P blocks and B blocks, motion-compensation consists of taking the inverse DCT output and adding it to the reference block(s) pixel values; reference block address is given by the motion-vector information decoded at the macroblock layer.
7. Finally, the image domain data is displayed. The display step includes color conversion from the YCbCr color space to the RGB space. Since Cb and Cr pixel data is half the resolution of

the Y data, upsampling needs to be performed during or prior to the YCbCr to RGB conversion phase. Additional upsampling of the pixel data may be required for display, e.g. the player might have to display the image in a larger window than its original resolution.

Steps 2-7 are compute-intensive and are the main bottlenecks to real-time MPEG1 video playback for a software based video player. In a practical implementation, some form of error-concealment must also be employed during video decoding. In the next section, we describe some of the optimizations incorporated in HP's MPEG1 video player.

3 Algorithm and Architectureal Enhancements

3.1 Enhancement Methodology

The basic approach was to examine the workload associated with each step of the decoding process outlined in the previous section and then develop algorithms for some of these steps that would lead to a reduced workload. The performance goal was to get a 10 - 15 fps playback of SIF resolution (352 x 240) MPEG1 compressed video and audio assuming that all of the enhancements were restricted to the algorithm level only.

A simple analysis of the video decoding steps outlined in the previous section indicated that the bulk of the execution time was spent in the IDCT step (46.4%) followed by the Display step and then the Motion-compensation step. Other steps in the decoding process consumed negligible time. Thus, algorithm and architectural enhancements were primarily targeted at these steps of the video decoding process.

3.2 Video Decompression - IDCT Optimization

In MPEG1 compressed video, an analysis was performed on the bitstreams. It was observed from this analysis that the IDCT computations were often performed on sparse matrices. Thus, if one could determine the nature of this sparseness, one could reduce the computation load of the IDCT. In order to determine the sparseness without additional overhead, it was found that by viewing the huffman decoder, inverse-quantization and the IDCT computations as a single system, it is possible to develop a computation procedure that reduced the workload for these three steps combined. This is the approach that is adopted in HP's MPEG1 player.

Inverse quantization can be performed within the huffman decoder, thereby, reducing accessing the same data twice. A low complexity IDCT algorithm was developed; its worst case performance is 80 multiplies and 464 additions for a 8x8 block. By exploiting the sparseness information, this IDCT algorithm yields an average performance of 46 multiplies and 253 additions for a 8x8 block. A lookup table based approach can be used for the multiply operation since the constants used in the IDCT were relatively few. Lookup table accesses are memory accesses which may be time-consuming. Instead, in the IDCT, the constants were chosen such that the multiply operation can be performed with a minimum number of shift-and-add operations and yet maintain good accuracy within the IDCT. The shift-and-add operation was

DOCKET ALARM

Explore Litigation Insights

Docket Alarm provides insights to develop a more informed litigation strategy and the peace of mind of knowing you're on top of things.

Real-Time Litigation Alerts

Keep your litigation team up-to-date with **real-time alerts** and advanced team management tools built for the enterprise, all while greatly reducing PACER spend.

Our comprehensive service means we can handle Federal, State, and Administrative courts across the country.

Advanced Docket Research

With over 230 million records, Docket Alarm's cloud-native docket research platform finds what other services can't. Coverage includes Federal, State, plus PTAB, TTAB, ITC and NLRB decisions, all in one place.

Identify arguments that have been successful in the past with full text, pinpoint searching. Link to case law cited within any court document via Fastcase.

Analytics At Your Fingertips

Learn what happened the last time a particular judge, opposing counsel or company faced cases similar to yours.

Advanced out-of-the-box PTAB and TTAB analytics are always at your fingertips.

API

Docket Alarm offers a powerful API (application programming interface) to developers that want to integrate case filings into their apps.

LAW FIRMS

Build custom dashboards for your attorneys and clients with live data direct from the court.

Automate many repetitive legal tasks like conflict checks, document management, and marketing.

FINANCIAL INSTITUTIONS

Litigation and bankruptcy checks for companies and debtors.

E-DISCOVERY AND LEGAL VENDORS

Sync your system to PACER to automate legal marketing.