Modeling and Design of Video Codecs Using Ptolemy

Jong-il Kim

The author is with the Laboratory for Image and Video Engineering, Department of Electrical and Computer Engineering, The University of Texas at Austin, Austin, TX 78712-1084, USA. E-mail: {jikim}@vision.ece.utexas.edu. Web: http://anchovy.ece.utexas.edu/{ jikim}.

Keywords

Embedded Systems, MPEG, H.263, H.263+, Multimedia, Ptolemy

Abstract

The key contribution of this project is a realization of modular, scalable and extensible simulation and design paradigm for video codec based on the system-level modeling and its implementation on extensible framework. To support this new simulation and design paradigm, generic dataflow of video codec system is modeled by homogeneous synchronous dataflow (HSDF) and implemented on the heterogeneous framework such as Ptolemy.

As a result of HSDF modeling, generic video codec is decomposed into ten basic modules which are fully composable. Since the design paradigm is modular and extensible, any video codec standards (e.g., H.263, MPEG-4) can be mapped on that basis targeting for various hardware architecture including DSP and ASIP.

The developed video codec libraries are dynamically linked to Ptolemy for simulation, and once the codec parameter is fixed during simulation level those libraries can be statistically linked for permanent use for design or retargeting level. The models of computation and the proposed design paradigm of video codec algorithm also helps automate partitioning of the design into hardware and software.

I. Introduction

Multimedia industries are witnessing a rapid evolution toward integrating complete systems on a single chip. The most important characteristic of multimedia applications is the rapid growth of the complexity of the design. As an example, the new videophone and teleconferencing standard (H.263+) [1] is much more complex than the previous H.263 and H.261 standards [2], and the continued evolution of video coding standards from MPEG-1 and MPEG-2 to MPEG-4 [3] is accompanied by a significant complexity increase with user flexibility. The conventional approach for system-level design and integration of video processing technology, combined with the level complexity and extremely short product development cycles, cannot find robust, cost-effective solutions simply by implementing large parts of the system functionality in software running on application-specific instruction set processor (ASIP) cores.

To support simulation and its extension to synthesis for video encoder/decoders (codecs) as new design paradigm, this project focuses on the system-level modeling of video processing systems, and its implementation on retargetable framework. Based on the model, new video compression and decompression standards (e.g., H.263, MPEG-4) is implemented in a formal, consistent and extensible framework such as in Ptolemy with well-defined and optimized processing primitives. The models of computation and an implementation of image and video processing algorithm will also help automate partitioning of the design into hardware and software [4]. The proposed design scheme is implemented and validated in Ptolemy [5].

The following section II summarizes video compression standards and their design perspectives using embedded processors. Section III describes a new design paradigm and implementation. Section IV concludes this project.

II. Background

The primary features of the standards for multimedia applications such as video codecs and teleconferencing are the significantly increasing complexity and fast emerging. At the time of MPEG-1 chips were available at the market, MPEG-2 standard was fixed, and before the commercial market for the previous H.263 standard [2] begins, new videophone and teleconferencing standard (H.263+) [1] emerges with multiple options and increased complexity. The continued evolution of video coding standards [3] from MPEG-1 and MPEG-2 to MPEG-4 is accompanied by significant complexity increase with object-oriented composition by user. The multimedia video codec standards recommended by the Motion Picture Experts Group (MPEG) and International Telecommunication Union (ITU), and their embedded design approaches are discussed in this section.

A. Overview of Video codec Standards

From a technical point of view, video compression standards can be categorized into two groups. One group is for low-complexity, frame-based coding schemes including MPEG-2 and H.263. The other one is for high-complexity, object-based coding algorithms such as MPEG-4, H.263+ [3], [6], [2]. In MPEG-2 and H.263, the video information is assumed to be rectangular, or of fixed size, displayed at a fixed interval. With predecessors like MPEG-1 and H.261, MPEG-2 and H.263 standards have been applied to CD-ROM, Digital Video Broadcast (DVB), Digital Video Disk (DVD), teleconferencing and High Definition TV (HDTV).

In MPEG-4, the concept of a video object (VO), and video object plane (VOP) have been introduced [3]. VO and VOP correspond to entities in the bitstreams that a user can access and manipulate (with, e.g., cut and paste operations). The VOP can have arbitrary shape. At the decoder side, composition information is sent to indicate where and when each VOP is to be displayed along with the VOP itself. Also, the user at the decoder

side may be allowed to change the composition of the scene displayed by interacting on the composition information. H.263 and H.263+ can be applied for object-oriented coding with additional information such as binary mask [2], but the functionality is not so broad as that of MPEG-4 [6].

MPEG-4 core coder has a similar structure to H.263 and MPEG-2, and MPEG-4 generic coder requires shape information to support arbitrary shape VOP. The overall MPEG-4 encoder blocks composed of Discrete Cosine Transform (DCT) and quantization which performs to compress the input VOP utilizing spatial correlation of the texture, and the inverse operations (IDCT and inverse quantization) reconstruct the input VOP to frame store block for motion estimation. Next input VOP is compared with the best motion estimated block, and differential data between the input and predicted block is coded if it generates more compact code. Decoder block is embedded in encoder because of differential coding.

H.263+ is an extension of H.263 [1], [2], providing twelve new negotiable modes. These modes improve compression performance, allow the use of scalable bitstreams, and the user flexibility. The basic coding techniques are common to other standards like MPEG standards, and H.263-based coder is bitstream compatible with MPEG-4. As described in this section, the functionality of MPEG-4 coder includes other coding standards, and it will provide generic technology for multimedia communication applications and services.

B. Embedded Processors for video codecs

Table I gives a list of products in the areas of MPEG and videophone multimedia applications. This category of the multimedia processors is of particular interest for this project as a modular design approach, since many common blocks like MPEG-1, MPEG-2, H.263, MPEG-4 video encoders and decoders are shared over a variety of emerging applications like set-top boxes for satellite and cable digital TV, DVB, DVD and PC-based multimedia

accelerators [7].

As Table I shows that virtually all of the MPEG codecs are based on application-specific instruction set processor (ASIP) cores and very long instruction word (VLIW), or on reduced instruction set processor (RISC) and custom hard-wired logic. The widespread decision to develop special ASIP cores highlights two important aspects of the multimedia market. First, cost is a critical factor. If a tailored architecture can deliver the same function at a low cost, then it will become a viable engineering choice. Second, none of the current 32-bit microprocessor architectures is well-suited to the task performing routines necessary for video compression and decompression, e.g., motion estimation, DCT and IDCT.

Two key features of embedded processors for multimedia applications are the increasing diversity of processor architectures, driven by low-cost consumer-oriented markets, and the wide range of architecture partitioning strategies and the diversity of building blocks. Many products performing the same function are designed using very different combinations of RISC, ASIP, and/or hard-wired co-processors.

But the significant growth of application complexity and the fast emerging new standards require new design strategies. As a results, many functions currently in hardware can be performed in software, and an ASIP will be necessary only for performance or cost reasons. However, because of the heterogeneity in video processing systems, no one design method is applicable to the entire system.

III. New Design Paradigm and its implementation

Observing the fast emerging multimedia standards and significantly increasing complexity, new design paradigm is suggested for the system level designers. While system designers seek to shorten design cycles, improve quality, and reduce cost of the systems,

TABLE I

EMBEDDED PROCESSORS FOR MULTIMEDIA

Company	*Processor*	*Applications*
Philips	VLIW DSP	MPEG-1/2 Decoder, H.261
NTT	Two RISC hard-wired ASIP	MPEG-2 Encoder
TI	MCU/DSP ASIP	MPEG-2 Decoder
TCEC	VLIW hard-wired DSP	MPEG-2 Decoder
SGC-Thomson	VLIW DSP ASIP	H261 video phone

no single design method is applicable to the entire system such as multimedia in that they contain a variety of algorithms implemented by many software and hardware technologies. Video processing and conferencing systems area heterogeneous in that they contains a variety of algorithms implemented by many software and hardware technologies, so that no one design method is applicable to the entire system design. Successful system design approaches integrate application-specific design methods and implementation technologies in a formal, consistent, extensible framework. The entire system can be specified, simulated, and synthesized in a formal, consistent framework, and extensibility supports new algorithms, tools, languages, and architectures to be integrated into the given framework to be rapidly retargeted to new technologies.

To support this new simulation and design paradigm, generic dataflow of video codec system is modeled by homogeneous synchronous dataflow (HSDF) and implemented on the heterogeneous framework such as Ptolemy. The model of computation is a key to the heterogeneous approach, which is the semantics of the interaction between modules or components in domain [8] [9]. The design approach should be based on the use of one or more formal models to describe the behavior of the system at a high level of abstraction.

During the implementation process, many different specifications and modeling techniques can be used [4],[11],[10]. HSDF model is composable because the number of token produced and comsumed on each arc is one, so that HSDF model is adequate in simulation level which requires dynamic composing of the HSDF libraries and modification of the algorithm.

For the purpose of overall video codec system design purpose, detail modeling for the video codec such as MPEG-4 and H.263 cannot be fixed but variable. If the video codec combines with some network or communication protocols, or if it involves multiple object scalability, DE model is better for exact simulation because of its additional time tag semantics. Ptolemy gives rich libraries for SDF and allows mixed model simulation and supports code generation in C and VHDL. Starting from the developed SDF libraries for generic video codec, mixed-domain simulation can be possible. The overall HSDF modules for generic video coded is depicted in Figure 1, and Figure 2 shows an example of stars generation as an applicatin for motion estimation and compensation. Once the SDF model is fixed, the construction of efficient loop structure from SDF graphs allows the advantages of inline code generation under stringent memory constraints. A variety of efficient design metric exist, so that during the modeling and optimizing the entire systems, different metric will be applied.

IV. Conclusion

The fast emerging standards and significantly increasing complexity require a new design paradigm. To support this new simulation and design paradigm, generic dataflow of video codec system is modeled by homogeneous synchronous dataflow (HSDF) and implemented on the heterogeneous framework such as Ptolemy.

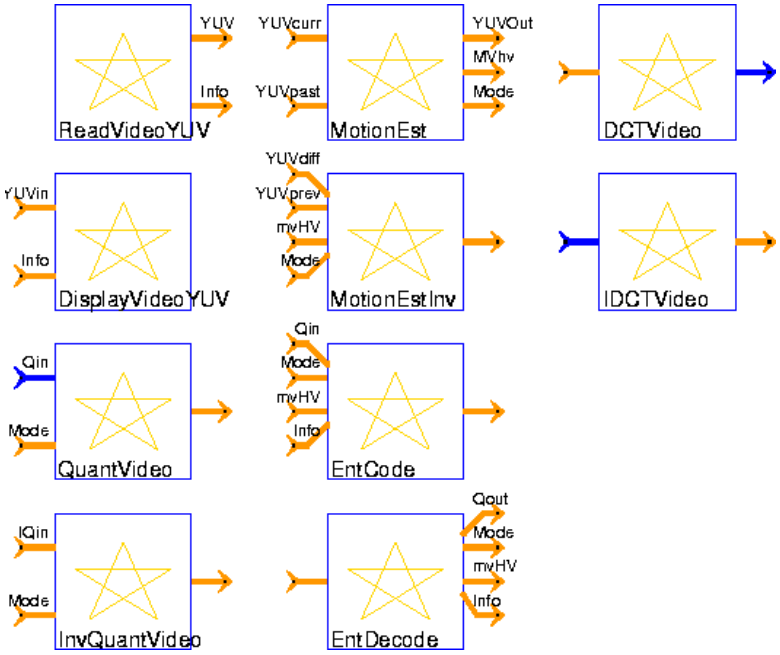

Fig. 1. Homogeneous synchronous dataflow modules for generic video codec

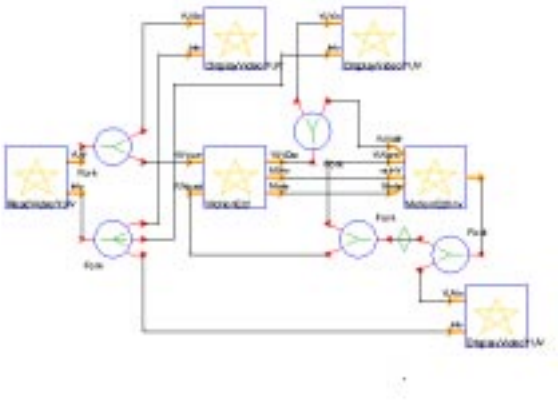

Fig. 2. Implementation of motion estimation and compensation system using HSDF modules

The developed video codec libraries are dynamically linked to Ptolemy for simulation, and once the codec parameter is fixed during simulation level those libraries can be statistically linked for permanent use for design or retargeting level. The models of computation and the proposed design paradigm of video codec algorithm also helps automate partitioning of the design into hardware and software.

References

[1] G. Côté, B. Erol, M. Gallant, and F. Kossentini, "H.263+: video coding at low bit rates," *submitted to IEEE Trans. on Circuits and Systems for Video Tech.*, http://www.ece.ubc.ca/image/papers/h263plus.pdf.

[2] J. Hartung, A. Jacquin, J. Pawlyk, J. Rosenberg, H. Okada, and P. E. Crouch, "Object-oriented H.263 compatible video coding platform for conferencing applications," *IEEE Journal on Selected Areas in Communications*, vol. 16, no. 1, pp. 42–55, Jan. 1998.

[3] T. Sikora, "MPEG digital video-coding standards," *IEEE Signal Processing Magazine*, vol. 14, no. 5, pp. 82–100, Sept. 1997.

[4] J. Pino, M. C. Williamson, and E. A. Lee, "Interface synthesis in heterogeneous system-level dsp design tools," in *Proc. IEEE Int. Conf. Acoust., Speech, and Signal Processing*, Atlanta, GA, May 1996.

[5] S. S. Bhattacharyya, P. K. Murthy, and E. A. Lee, "Synthesis of embedded software from synchronous dataflow specifications," *to appear in Journal of VLSI Signal Processing*, http://ptolemy.eecs.berkeley.edu/papers/publications.html/pubs_1998.html.

[6] L. Chiariglione, "MPEG and multimedia communications," *IEEE Trans. on Circuits and Systems for Video Tech.*, vol. 7, no. 1, pp. 5–18, Feb. 1997.

[7] P. Paulin, C. Liem, M. Cornero, F. Nacaba, and G. Gossens, "Embedded software in real-time signal processing systems: Application and architecture trends," *Proc. of the IEEE*, vol. 85, no. 3, pp. 419–435, Mar. 1997.

[8] W. T. Chang, S. Ha, and E. A. Lee, "Heterogeneous simulation -mixing discrete-event models with dataflow," *Journal on VLSI Signal Processing*, vol. 13, no. 1, Jan. 1997.

[9] E. A. Lee and A. Sangiovanni-Vincentelli, "A framework for comparing models of computation," *Revised from Memorandum UCB/ERL M97/11*, Jan. 1997.

[10] J. Liu, M. Lajolo, and A. Sangiovanni-Vincentelli, "Software timing analysis using HW/SW cosimulation and instruction set simulatior," in *6th International Workshop on Hardware/Software Co-design*, Seattle, Washington, Mar. 1998.

[11] M. C. Williamson and E. A. Lee, "Synthesis of parallel hardware implementation from synchronous dataflow graph specifications," in *Proc. IEEE Asilomar Conf. on Signals, Systems, and Computers*, Nov. 1996.