An FIR-Based Video Format Conversion IC Design

Yue Lu, Jun Sun, Dong Wang and Zhigang Wang

Shanghai Jiao Tong University
Institute of Image Communication and Information Processing
1954 Hua Shan Road, Shanghai 200030, China
{yuelu, sunjun, wangdong, zgwang}@cdtv.org.cn

Abstract

High quality and versatile video format converters play an indispensable role in the emerging DTV age. This paper presents our design of an IC capable of cross-converting between various DTV standards. A multi-phase FIR-based filtering algorithm is developed to perform the video scaling tasks. Compared with other methods, this approach is relatively easy for hardware implementation and leads to superior robustness and fairly good output qualities. In the paper, we explain the whole architecture of our IC and some of its key components. The design is a suitable and high quality solution to consumer applications. Meanwhile, it can also serve as a building block for the more advanced systems with a motion-compensated de-interlacing stage.

1. Introduction

The advent of the Digital Television (DTV) [1] age is doubtlessly a major revolution in the realm of consumer electronics. Thanks to the recent advancement in source coding and channel transmission technologies, HDTV can bring superior-quality visual and audio content to everyone's living room. However, the huge cost to set up a new DTV station, the lack of programs in HDTV format, and the numerous amount of incompatible analog television sets people already own at home …… all of them make the transition to the DTV age a step-by-step process. It is foreseeable that HDTV, SDTV and even traditional analog TV might coexist for a long period of time. In this case, materials produced in one format must be up or down-converted before they can be properly transmitted and received in a system adopting another format.

Modern CRTs become increasingly larger and brighter. Unfortunately, the frequency response of the human eyes is extended on bright sources, and this makes field-rate flicker more visible and objectionable. One solution to this problem is to double (or at least increase) the field rate. The flicker is then beyond the response of the eye. A format converter can be used here to perform this task.

Another reason to introduce a video format converter is the recent trends in using liquid crystal displays (LCDs). Unlike the traditional CRTs that can accommodate many different raster formats, each LCD has its 'native' resolution, i.e. the number and structure of the pixels that it can display is pre-determined. In this case a format converter should be used to 'translate' different input sources to the LCD's 'native' resolution.

From the above discussion, we can see that high quality and versatile video format converter is an indispensable component in the emerging DTV age. To meet this demand, we propose a DTV format conversion IC that has the following key features:

- Input and output formats capable of up to 1920 × 1080 pixel resolution.
- Arbitrary conversion ratio between input and output spatial sampling rate.
- 3/2 pull-down can be removed for true frame processing of film material.
- 50Hz to 75Hz or 60Hz to 90Hz frame/field rate up-conversion.
- Local pan, tilt and zoom of a specified portion of the image
- Picture-in-Picture (PIP) mode supported
- System parameters, including the filter coefficients, can be changed at run time via an I^2C interface.

The rest of the paper is organized as follows: In section 2 we discuss the algorithms used in this chip. In section 3 the whole architecture of our IC and some of its key components are presented and explained. In section 4 we describe the verification system used to test the chip. Section 5 provides a summary and an outlook to future work.

2. Algorithm Overview

2.1. FIR-based Scaling

Theoretically, video format conversion (VFC) can be viewed as a spatial-temporal sampling lattice conversion problem [2] [3]. The state-of-art of VFC techniques [4] [5] [6] exploits multi-frame information, i.e. use motion compensation (MC). These MC-based methods have been shown to outperform other methods in terms of resulting image quality. However, they have several major drawbacks. First of all, these methods need expensive motion estimation between consecutive frames, which will greatly complicate the whole system and increase the cost of the IC. Secondly, their performances rely heavily on the precision and correctness of the estimated motion vectors. Since wrong motion vectors can give rise to very objectionable artifacts in the output images, all the MC-based methods have to take careful measures to increase their robustness against incorrect motion vectors [7]. This will again increase the complexity of the whole design. For the cost-sensitive consumer applications, those drawbacks cannot be neglected.

In our design, we have turned to an intra-field FIR filtering algorithm to perform the VFC. Strictly speaking, intra-field filtering assumes that the input video is a band-limited signal while this assumption hardly holds in real world. However, as shown by the experiments, neglecting this assumption violation and performing intra-field filtering anyway still result in fairly satisfying subjective image quality. Although we have observed some small artifacts such as ringing and blurring, they are still within the acceptable range and the overall quality meet the needs of consumer market quite well. Another advantage of the filter-based methods is its superb robustness. The performance degradation from these methods is characterized as global and gradual and is less objectionable to the viewers than are local and abrupt artifacts caused by incorrect motion vectors.

2.2. Interpolation Kernel

The idea of the FIR-based scaling is quite simple and reference [8] gives the technical details of it. The difficulty here lies in the proper design of an interpolation kernel. In order to mitigate the ringing artifact caused by the direct truncation of the sinc function, we employ a smooth windowing function with finite support. Then, the interpolation kernel can be described as:

$$h(x) = W(x)\frac{\sin(x)}{x} \tag{1}$$

where $h(x)$ is the interpolation kernel and $W(x)$ is the windowing function.

There are many windowing functions that can be used. Pass-band and transition-band are two important factors for a specific windowing function. Generally speaking, a narrow pass-band will lead to image blurring and a wide transition-band will lead to aliasing. On the other hand, a narrow transition-band can cause ringing artifacts. Therefore there's a trade-off between these two factors. After many experiments, we finally choose the raised-cosine windowing function, which is defined as

$$W(x) = \frac{\cos(\pi\beta x)}{1-(2\beta x)^2} \tag{2}$$

The β in the above expression is called roll-off parameter and can be used to control the pass-band width and transition-band width.

In practice, we have also used several techniques including multi-phase filtering in our design. These methods can greatly reduce the computational complexity of the filtering process. In section 3 we will give a detailed account of them.

2.3. Film Material Pre-Processing

For materials originally produced in film format, they have a progressive content with a 'native' frame rate of 24Hz or 25Hz. In order to adapt to the existing PAL or NTSC format, either a 2:2 or a 3:2 pull-down is applied. Take the 3:2 pull-down for example, odd frames are split into two fields while even frames result in three fields. In this case, consecutive odd and even fields in a 60Hz 3:2 pull-down video may originally come from one progressive frame. If we first merge these two fields and perform FIR scaling on the progressive frame, the output image quality can be improved a lot. The reason for the performance gain is quite straightforward: the input information to the FIR-based scaling is doubled and hence, the requirement for band-limited signal is better fulfilled.

In our design, the fields merge operation for film materials are performed by a special scheduling unit in the input control module, which determines the "life cycle" of each field in the memory buffer.

3. Proposed Architecture

3.1. Architecture Overview

The architecture of our IC can be divided into five major components: input control module, FIR filter module, output control module, SDRAM controller, and parameter programming interface. Figure. 1 is the functional block diagram of the IC. There are two processing channels in this IC: master channel and slave

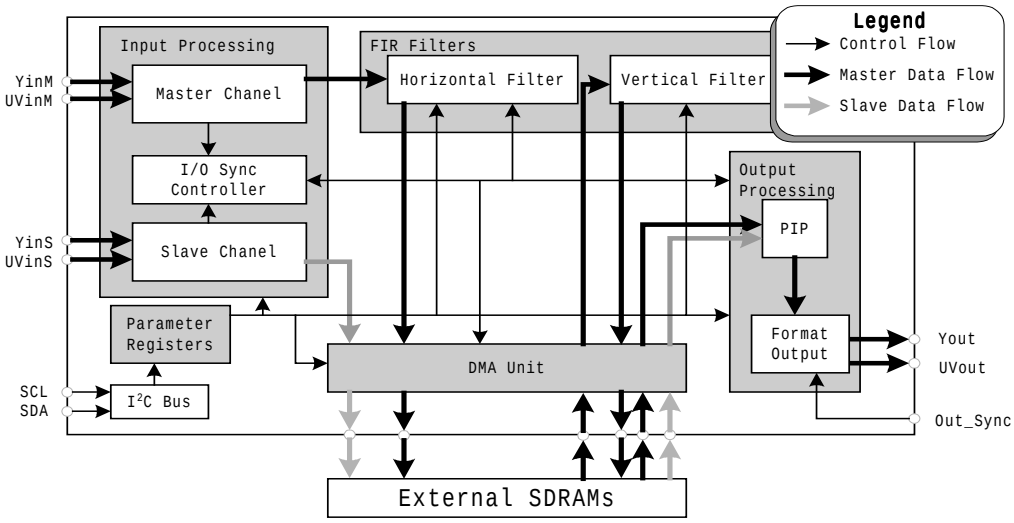

Figure. 1 **Functional block diagram of the IC**

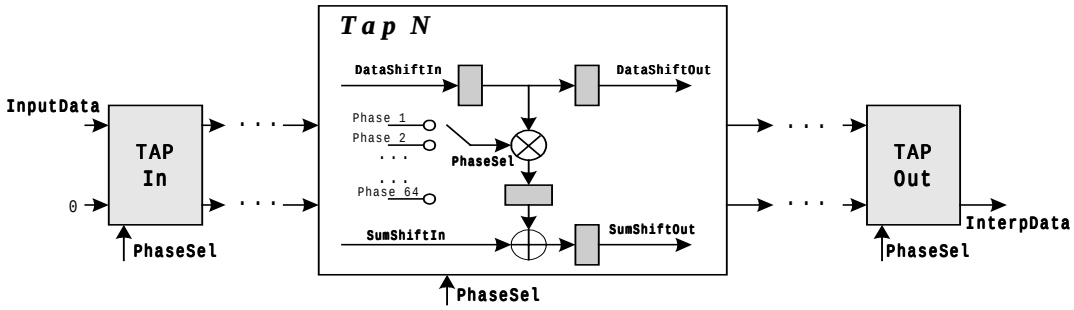

Figure. 2 **A multi-phase interpolation filter bank**

channel. The image data in the former channel undergo FIR-based scaling, while any input to the latter channel goes directly to the memory control unit and serves as the background in the picture-in-picture (PIP) mode.

3.2. Input Processing

The input control unit accepts two channels of 8/10-bit YCbCr 4:2:2 input signals along with their associated timing information. The input video to the master channel will be de-noised by a cost-efficient processing unit, which consists of an edge-extraction high pass filter and an edge-adaptive low-pass de-noise filter [9]. This combination will avoid the undesirable blurring artifact, which is usually a by-product of the low-pass de-noising operations.

Meanwhile, the input control unit is also responsible for the global synchronization tasks. In each channel, a simple four-state finite state machine (FSM) is designed to extract from the input streams their TRS (Timing Reference Signal) information, which indicates the starting or ending-point of frames/fields or lines. In turn, the extracted timing information is used to generate the control signals to synchronize the whole IC.

Film material pre-processing and frame/field rate conversion are also performed in the input module by a scheduling unit, whose output signals notify the SDRAM module when to hold or discard certain frames/fields.

3.3. FIR Filter

The FIR filter block is the heart of the whole IC. The principle of the FIR-based filtering is to weight the input samples with appropriate coefficients taken from the interpolation kernel. The value of the coefficient is determined by the phase (position lying between two input samples) of an output sample. Due to the requirement of our IC, the phases of the output signal can take arbitrary values in the legal range. Therefore, it is not viable for us to store all the possible coefficients set in the memory. Our solution to this problem is to store a large number of phases (say 128) in the on-chip memory. All the real output samples are rounded to the

nearest stored phase. Experiments have shown that this quantization method is both practical and efficient, and that the performance lost is hardly observable.

The multi-phase interpolation filters are implemented in a systolic structure, as shown in Figure. 2. All taps are cascaded and synchronized by a common clock signal. The PhaseSel signal is used to select the corresponding coefficients for a certain phase. Since the number of valid data at the input port is not equal to that at the output port, the input/output decimation/hold signals are used to specify whether the current output data is valid or not, and whether the interpolation filter's data registers need shifting or not. All those signals are generated by a field/frame re-constructor, which is composed of several counters controlled by the input unit. To further simplify the filters, we have also employed the PRF method described in [10]. After 3-route paralleling, algorithm reducing and architecture folding, about 60% of the interpolation filter's area can be saved.

3.4. SDRAM Controller

As an economical high-speed/density storage device, the SDRAM is chosen to be the external frame/field memory of the IC. In our system, three frame/field buffers are required: one slave channel buffer, one master channel horizontal processing buffer, and one master channel vertical processing buffer. While running in the HDTV mode, the IC requires a high throughput rate (up to 74.25×2 Mbytes/s for I/O). Thus, it is not practical to store all the buffered frames in a single SDRAM chip. In our design, we have separated the read and write operations into different SDRAM chips managed by a Ping-Pang scheme.

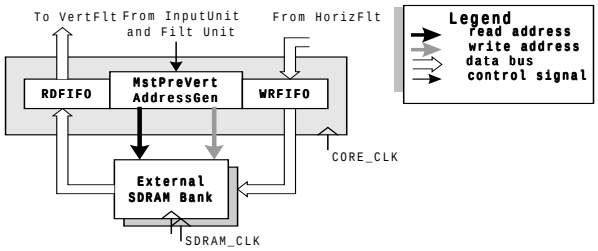

Figure. 3 **Master channel horizontal processing frame buffer controller unit**

Figure. 3 demonstrates the design of the master channel horizontal processing frame buffer controller. The other two SDRAM controllers have similar structures. The external SDRAMs run at a speed of about 100 MHz and the core runs at about 75MHz. The read or write FIFO provides a safe channel for data flow spanning different clock domain. Besides, the I/O FIFOs can also ensure the continuous stream from/to core despite the discontinuous access to the SDRAMs.

The AddressGen unit in the master channel generates the SDRAM access address. Video signals in the master channel will be transposed in SDRAMs, i.e. they will be read in horizontal or vertical order while written in the reversed order. However, directly storing frames in horizontal or vertcal order will result in misbalance between SDRAMs' input and output. Consequently an unreasonably large amount of FIFOs are required for data continuity. To solve this problem, a dedicated SDRAM data distribution algorithm is implemented in the AddressGen unit, which re-organizes the frame data in an interleaved order. This will greatly alleviate the demand for large on-chip FIFOs.

3.5. Output Processing

The output unit in the IC mainly performs two functions: PIP and output formatting. Prior to PIP processing, the video streams in the master and slave channels will be synchronized by the input unit. Meanwhile, the TRS signals will be inserted at proper positions according to the SMPTE standards.

3.6. Parameter Programming Interface

As the industrial standard interface for consumer ICs, I^2C has been well supported and can be conveniently implemented. Thus, we adopt I^2C (working in slave mode) as the communication interface for our IC.

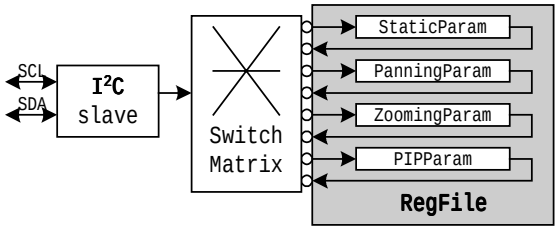

Figure. 4 **Parameter programming unit**

All the parameters in this IC can be grouped into two categories: the static parameters and dynamic parameters. The former mainly concerns with system configurations such as SDRAM parameters, display resolution and are only initialized once after powering on; the latter includes those parameters about zooming, panning and PIP settings etc, and can be changed at run-time. All these parameters are stored in an on-chip register file, which is implemented by shift registers to simplify address-decoding circuits. Furthermore, we have segmented all these shift registers into several sub-groups according to the similarities in their functions

(see Figure. 4). Thus, the parameters can be set much more selectively and efficiently.

4. Design Verification

Currently we have finished the RTL level design of this chip. Since the verification of our design via pure software simulation is a very time-consuming process, we have resorted to FPGA emulation in the verification stage. Figure. 5 shows the block diagram of the verification system. The board can receive real-time digital video streams through the PDI interface and half their frame rate via frame buffers. This slowdown of the input signals is necessary, since the FPGA chip cannot run under the desired frequency. The FPGA emulates the behaviors of our design and sends out the processed result. One channel of the result can be directly received by the computer and stored in disks for analysis, while video streams in the other channel are recovered to the original sampling rate and displayed on the monitors. The key characteristics of the above design are listed in Table. 1.

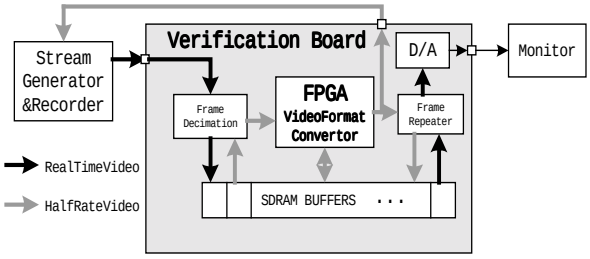

Figure. 5 **The verification system**

Table. 1 **Key characteristics of the FPGA implementation**

Device	Altera Apex20K1000-1
Package	BGA 652-Pin
LE	35,342
ESB	44 Segments
Max Frequency	38 MHz

5. Summaries and Future Work

In this paper we have presented our design of a video format conversion IC. We have proposed a multi-phase FIR-based filtering algorithm to do the video up and down-conversion tasks. Compared with other more advanced approaches, the FIR-based approach is relatively easy for hardware implementation and leads to superior robustness and fairly good output image qualities. The whole architecture of the IC and some of its key components are explained in the paper. Besides, we have verified our design through an FPGA emulation system.

The current state-of-art video format conversion algorithms exploit the motion information between consecutive fields. However, cost-efficiency and robustness are two main challenges to those methods. One alternative way is to only estimate and compensate the global motion in the image sequences. Since camera motion and operation are very common scenarios in a video sequence, the global motion-based algorithms share the same advantages of the MC-based approaches in many situations. However, since the global motion information is estimated from the entire image, it usually can be obtained with greater robustness and precision, and often with lower costs. Therefore, a video format conversion IC equipped with global motion compensation is a promising field for further research. Currently, we are developing our second IC along this direction.

References

[1] ATSC, "ATSC Digital Television Standards (Rev A)," Doc.A/53A, Apr.2001.

[2] A.M. Tekalp, "Digital Video Processing," Prentice Hall, New Jersey, 1995.

[3] E. Dubois, "*The Sampling and Reconstruction of Time-Varying Imagery with Applications in Video Systems*," *Proc. IEEE*, 73(4), 502-522 (1985).

[4] G. de Haan and E.B. Bellers, "*Deinterlaing - an overview*," *Proc. IEEE*, 86(9), 1839-1857 (1998).

[5] A. J. Patti, M. I. Sezan and A. M. Tekalp, "*Digital Video Standards Conversion in the Presence of Accelerated Motion*," *Signal Process. Image Commun.*, vol.6, 213-227 (1994).

[6] L. Vandendorpe, L. Cuvelier, etc., "*Motion-Compensated Conversion from Interlaced to Progressive Formats*," *Signal Process: Image Commun.*, vol.6, 193-211 (1994).

[7] O.A. Ojo and G. de Haan, "*Robust Motion-Compensated Video Upconversion*," *IEEE Trans. Consumer Electronics*, 43(4), 1045-1056 (1997).

[8] A.V. Oppenheim, R.W. Shafer, "Discrete-Time Signal Processing," Prentice-Hall, New Jersey.

[9] K. Jostschulte, A. Amer, etc., "*Perception Adaptive Temporal TV-Noise Reduction Using Contour Preserving Prefilter Techniques*," *IEEE Transactions on Consumer Electronics*, 44(3), 1091-1096 (1998).

[10] D. Wang, W. Wang, etc., "*A Low Area/Power FIR Design Based on PRF Framework*," Technical Report HDIC01-15, Institute of Image Communication and Information Processing, Shanghai Jiao Tong University.