Design of Audio Embedding and De-embedding for 3G-SDI Based on FPGA

Yaqing Liu^{1,*}, Lunhui Deng²

¹Communication University of China, Department of Computer Science, Beijing, China

²Communication University of China, Department of Computer Science, Beijing, China

Abstract. This design introduces the theoretical basis of digital audio embedding and de-embedding, and proposes a solution that Verilog language can be used to achieve 3G-SDI audio embedding and de-embedding. SDI video and audio data are input to the FPGA, and the audio signals can be embedded in the SDI line blanking after processing. Moreover, some auxiliary information is embedded in the SDI data, when you need these auxiliary information, you need to use the audio de-embedding process. The process of audio de-embedding is inversed with the process of embedding. It has been proved through practice that this scheme can effectively embed digital audio in SDI data stream, synchronize audio and video data, and can de-embed audio signal. The design is very versatile and can improve the efficiency of the design, thus effectively reducing the cost of the product.

1 Introduction

3G-SDI signal has been widely used in broadcasting industry, with the continuous development of the security industry, its advantages of high speed and uncompressed numbers are gradually excavated. A large number of 3G-SDI series products have been launched on the market, including SDI conversion equipment, SDI digital switching matrix devices, and SDI distributor, etc. These devices use 3G signals and are compatible with 1.5G signals, which can be used for long distance transmission to meet the diverse needs of the wide range of users.

The broadcast of TV programs is usually accompanied by synchronous sound signals. It is difficult to synchronize the video and audio signals that are transmitted separately. Digital audio embedding means embedding the audio signal into the blank data area of the video signal, thus realizing the simultaneous transmission of video and audio signals on the same data bus. The process of audio De-embedding is opposite.

2 I2S interface specification

Audio can be divided into analog audio and digital audio, and analog audio can be converted into digital audio. I2S is a bus standard set by PHILIPS for data transmission between digital audio devices, such as CD player, digital sound processor and digital TV system. Its transmission of clock and data signals are independent. By this way, the clock signals and data can be separated, so that the distortion caused by time difference is avoided, which saves users the cost of professional equipment to resist audio jitter.

The I2S standard not only stipulates the hardware interface specification, but also specifies the format of digital audio data. It has three main signals: serial clock (SCLK), frame clock (LRCK), and serial data (SDATA). Sometimes, in order to synchronize the system better, we need to transmit another signal MCLK, called the main clock, also called the Sys Clock, which is 256 times or 384 times the sampling frequency[1].

I2S signal timing is shown in Fig. 1.

Fig. 1. Caption of the Figure 1. Below the figure.

No matter how many bits of data are available, the I2S format signals are always the first to transmit the highest data in the transmission. Therefore, the highest bit location is fixed, and the lowest bit position depends on the effective digit of data, that is, the effective digit of the receiver and the transmitter can be inconsistent[2].

If the word length of the system transmitter is larger than that of the receiver, the redundant low bit data in the data frame can be abandoned. If the receiver can handle more effective bits than the sender, it can make up the remaining bits by itself (usually 0 bits), which can effectively reduce the bit error rate in transmission and make the devices more convenient to connect with each other.

* Corresponding author: 15001207598@163.com

Fund code: 2015BAK05B01 ; 2017XNG1602

3 Audio embedding principle

Non-video data may be transmitted by 288 words of data in the active line of the SDI video. The data of each line in the filed blanking is the same, and is distinguished by "EAV" and "SAV".

"SAV" and "EAV" have the same first 3 data words, which have unique identities. By detecting these three data, we determine the starting point and end point of the auxiliary data area embedded in the audio packet[3].

Fig. 2. The format of auxiliary data.

The SMPTE425M interface standard defines the format of the auxiliary data area as shown in Fig. 3.

EAV	Ln	CRC	Auxiliary data	SAV
3FF 000 000 XYZ	Ln0 Ln1	YCR0 YCR1	Control	3FF 000 000 XYZ
3FF 000 000 XYZ	Ln0 Ln1	CCR0 CCR1	Data	3FF 000 000 XYZ

Fig. 3. The format of the auxiliary data area.

According to the SMPTE425M standard, the auxiliary data format is mainly made up of EAV, Ln, CRC, auxiliary data area and SAV. Line/field synchronization information and the state of vertical and horizontal hidden areas all exist in XYZ. 3FF, 000 and 000 are presynchronous codes, which are reserved for timing recognition[4]. They can accurately identify the start of synchronization information between SAV and EAV. The location of a row in a frame is represented by line number Ln. The role of check word CRC is to check the whole line video data.

When the audio is embedded, the left and right sound channels are mapped to three 10 bit audio words, respectively, X, X+1, and X+2. In the auxiliary data package, the head of the packet is made up of 6 state words: the flag of the auxiliary packet ADF(000,3FF,3FF), auxiliary packet type (DID), the data block number (DBN), and the data counter (DC)[5].

Based on the above theoretical analysis, the design proposes a program of audio embedding and de-embedding, hardware design shown in Fig. 4.

Fig. 4. Hardware design.

The external analog audio is converted into FPGA after the A/D conversion of the CS5340 chip. The 3G-SDI video flows through the LMH0344 chip, which is received by the SDI IP with the internal FPGA, and converted into a series of conversions. The 3G-SDI video stream is equalized by the LMH0344 chip, then inputs to the FPGA, and received by the SDI IP CORE. The audio and video signals enter the audio embedded module to embed the external audio signals into the 3G-SDI signal. Considering the analog audio, there are also the digital audio carried by the SDI signal. At this point, the embedded audio signal needs to be inlaid. Then the separated audio flows through CS4344 to be D/A converted and output. The 3G-SDI video stream is output via the LMH0303. At this point, the function of the whole system is achieved.

4 Hardware circuit design

The selection of chips is very important. Whether the chip's selection is correct will affect the development difficulty of the system hardware. If the choice is improper, it will directly lead to the failure to realize the function of the system and increase the design budget.

High speed serial signal will produce attenuation and distortion after long distance transmission, so cable equalization must be done first, then the regeneration serial differential signal is obtained. The LMH0344 3-Gbps SDI Adaptive Cable Equalizer is designed to equalize data transmitted over cable (or any media with similar dispersive loss characteristics). The equalizer operates over a wide range of data rates from 125 Mbps to 2.97 Gbps.

Fig. 5. SDI equalization circuit.

As shown in the figure above, the signal is entered into the circuit through the joint. Because the impedance of the coaxial cable is 75 Omega, the resistance of the selected 75 Omega is matched with the impedance. The main function of 1uF capacitor is to filter the input signal.

LMH0303 chip is used for output adaptive equalization.

Fig. 6. SDI equalization circuit.

The external input analog audio signals are directly input into the audio A/D converter CS5340.

CS5340 has the functions of sampling, A/D conversion and anti aliasing filtering, with digital filtering and simplification, and no external anti aliasing filter is required. The serial audio signal can not be played directly after the output of FPGA. It also needs to be converted into analog signal through the process of digital analog conversion.

Fig. 7. Audio A/D conversion circuit.

Analog electricity and digital electricity are separated. The analog power (VA) is 5V, the digital electricity (VD) is 5V, VD can be obtained from VL or VA. If we get it from VA, we must add a 51 ohm resistor in the middle. At this time, VD can only supply power for digital filter, and it can no longer supply power to other devices as a power source. The chip works in the main mode, the sampling clock LRCK and the serial clock SCLK are automatically obtained by MCLK.

Fig. 8. Audio D/A conversion circuit.

The design uses CS4344, which uses a linear analog low pass filter to realize the 4 order multibit Delta-Sigma modulator, which needs less circuit support.

CS4344 supports data format of standard audio sampling rate: 48, 44.1 and 32kHz under SSM mode, 96, 88.2 and 64kHz under DSM mode; 192, 176.4 and 128kHz under QSM mode. The audio data is input to the SDIN pin. The left and right clock (LRCK) determines the data channel, and the serial clock (SCLK) push the audio data into the input data buffer.

5 Implementation

The SDI video data enters the FPGA through the receiving module, and then the audio signal is embedded after the auxiliary information extraction and other operations. When the audio is embedded, the audio information can not be sent to the SDI video stream immediately. First, it should be stored. Because the number of packets embedded in the SDI attached data area is uncertain. When the ancillary data space of a row is coming, we can decide whether to embed the audio according to the embedded auxiliary information, and if the requirement is satisfied, the received data is embedded.

The audio embedding process is shown in Fig. 9.

Fig. 9. The audio embedding process.

The timing simulation waveform of the audio embedding module is shown in Fig. 10.

Fig. 10. The timing simulation waveform of the audio embedding module.

The simulation results show that, under the system clock, 24 bits of audio data are embedded in the attached interval of SDI to meet the design requirements.

Auxiliary information is embedded in the SDI attached data area, when these auxiliary information is needed, you need to use the audio de-embedding process. Audio de-embedding process shown in Fig. 11.

Fig. 11. Audio de-embedding process.

Fig. 12. The timing simulation waveform of the audio de-embedding module.

The location of audio data packet and audio control packet is different, the former is embedded into the chroma channel, and the latter is placed in the luminance channel, and it is processed separately when it is embedded, and they need to be respectively processed when de-embedding. In the process of audio de-embedding, we first need to detect the type of DID. If it is judged as an audio packet, it is necessary to detect the corresponding data block number (DBN), and then extract the audio channel information, and finally output the packet data after verification. If it is judged as the audio control packet, the corresponding sampling rate should be checked to verify the output of the channel after verification. At this point, the audio de-embedding is completed.

6 Summary

This paper proposes a scheme of audio embedding and de-embedding based on FPGA. Through the proof of software and practical application, this design can

achieve the function well. It has strong versatility and portability. With the continuous development of digital TV, meeting the continuous improvement of performance requirements will have strong core competitiveness.

References

1. Philips company. The I2S Bus specification version2.0 [S]. Holand : Philips Semiconductors,2000,(1):1-50.
2. Ji Xiaojun. FPGA based SDI audio coding and embedding design [D]. Nanjing Forestry University, 2016.
3. Ni Ning Ning. Audio embedded [J]. technology and application of radio and television technology, 2005 (11): 3-4.
4. Liu Fei. Design of FPGA based HD-SDI audio embedded system design [D]. Anhui: University of Science and Technology of China, 2011.
5. Qi na, Zou Feng.HD-SDI embedded audio data analysis [J]. radio and television information, 2014,12 (15): 2.
6. Zhang Jianfeng, Zhou Jianyang, Yin Huiqing. Design of audio inlay system in high definition digital video broadcasting [J]. Chinese integrated circuit, 2010 (3): 59-62.