A Hardware Framework for Smart Speaker Control of Home Audio Network

[1]Sneha Priya Sebastian, [2]Vineetha P Joseph

[1]Assistant Professor, [2] Assistant Professor
[1]Electrical and Electronics Engineering,
[1]Mar Baselios Christian College of Engineering and Technology, Kuttikkanam,
Kerala,India

***Abstract- In* this paper, an intelligent audio system controlling multiple speakers through a network cable and a USB cable used as an interface to a living room PC to connect to the proposed main controller. More distributed approach to multi-channel systems is. The main controller connects all speaker channels serially via a network cable through subcontrollers that are distributed across the rooms. The network cable interface between the sub-controllers and the main controller uses a custom-made transmission protocol. This protocol allows the main controller to provide multiple CD quality audio sources to different locations in a house. Furthermore demonstrate a bidirectional communication capability to allow seamless personalized listening at all locations in the home with the proposed protocol between the main controller and the sub-controllers. Multi-channel speaker systems configured as a home network appliance in modern houses are still mostly based on analog systems. These audio system cause connection and maintenance problem because of complex wiring. The main system consists of a USB 2.0 controller including the 8051-core and Altera's Cyclone II FPGA**.

I. Introduction

Advances in home networking have made high bandwidth digital access to all corners of a house possible. Consequently, personalized multi-channel audio is becoming a reality. For example, let's assume there are two persons with different musical tastes in the house doing various duties. If it is possible to track where an individual is, by using RFID or a similar device, the personalized audio can follow the individual. Not only does this save unnecessary power of activating all speakers, but it also allows the two individuals to listen to their own music. However, if the two persons are in the same room, the central controller has to be programmed to decide which sound stream should be played. Moreover, each of the speakers can communicate with the central controller to relay information about the room occupancy status. Unfortunately, multi-channel speaker systems configured as a home network appliance in modern houses are still mostly based on analog systems. These systems use a central multi-channel amplifier to drive each speaker and require a lot of wiring. Since speakers are devices which finally generate sounds for human beings, the number and the sound quality of speakers play an important role in the performance of audio systems. With the popularity of HTSs (Home Theater systems), 5.1-channel audio systems with 6 speakers have become common. In addition, 7.1-channel audio systems with better performance are going to be in the market soon. In the near future, 9.1-channel audio systems with 10 speakers may be developed if we consider the trend of current audio technology.

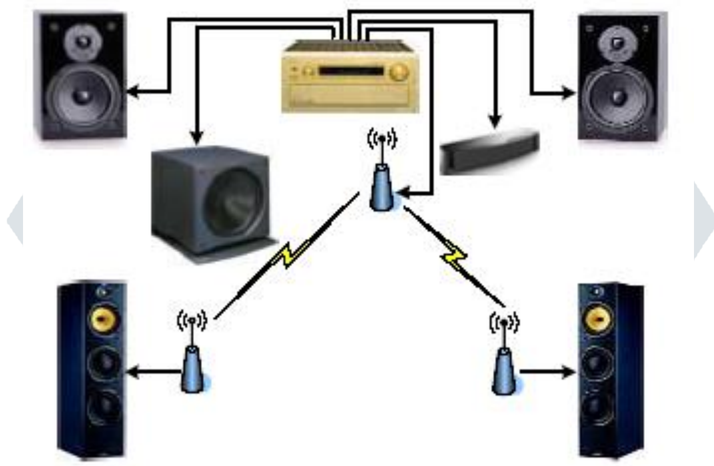
How to connect several speakers which belong to an audio unit is also very important if we consider the increasing number of speakers in multiple-channel audio systems. An existing audio system has one-to-one connection between each channel of the audio unit and its corresponding speaker. Compared to two speakers for a traditional audio system, a relatively large number of speakers for a multiple-channel audio system may cause connection and maintenance problem because of complex wiring. So a new speaker connection technique which connects all speakers in a serial fashion and sends digital audio signals to speakers sequentially was presented [2]. The connection technique converts analog audio signals to digital signals, sends them to corresponding speakers, and then converts them back to analog signals for making sounds. The technique is modeled in VHDL, and designed with FPGA chips as well as analog-digital converters. For the modeling and design, we need to consider the packet generation, the digital data compression, and the recovery from possible signal loss during transmission.

IDEAS WHICH GIVE WAY TO THIS PAPER

II. Connections of speakers for HTS

Let us investigate connection techniques of speakers for multi-channel audio systems. Figure a shows two ways of connecting speakers for a 5.1-channel audio system. Each channel of the audio system requires its own connection wire to send analog audio signals to its corresponding speaker. That is, a 5.1-channel audio system needs 6 wires for 6 speakers respectively.

a.1) Multiple-wired connection



a.2) Partially wireless connection

Fig a. Current connection techniques of speakers for a 5.1-channel audio system.

As the most popular connection, 6 speakers are placed and wired around an audio unit of a 5.1-channel audio system as shown in Figure a.(1). In this case, non-experts for audio systems may be faced with the complex wiring difficulties, the use of expensive connection wires for quality sound, and the adjustment of mutual impedance [4].To alleviate these problems, a partially wireless connection technique, based on either the Bluetooth technique or the 2.4 GHz RF technique, has been used to reduce the number of connection wires as shown in Figure a.(2). In this case, there are also several problems we need to consider. First, only two rear speakers are connected in a wireless fashion. Second, additional transmission and reception hardware modules are required. Third, the wireless speakers only work properly if they are within 10m of an audio unit. Finally, there might be a noise problem due to interference [6].

III. SERIAL CONNECTION TECHNIQUE OF SPEAKERS

To solve the above mentioned problems, a new technique which connects speakers in a serial fashion and sends digital audio signals to corresponding speakers sequentially[1] was developed Moovin Song, Ohkyun Kwon, Yunmo Chung, "A Serial Connection Technique of Speakers for Multi-channel Audio systems," *IEEE Transactions on Consumer Electronics*. Vol. 51, No.2, pp. 611-616, May 2005.

.

A. Serial Connections by Digital Transmission

In this technique, speakers can be connected to each other, regardless of the number and order of speakers, due to the transmission of digital audio signals. For example, several methods of serial connections can be presented. First of all, a single connection can be used to connect all speakers serially. Alternatively, two connections can be considered to connect speakers, one connection for speakers on the left side, the other one for speakers on the right side. In addition, it is easy to build an audio system with any number and configuration of speakers as shown in Figure b, which is an example of a connection for a 5.1-channel audio system.

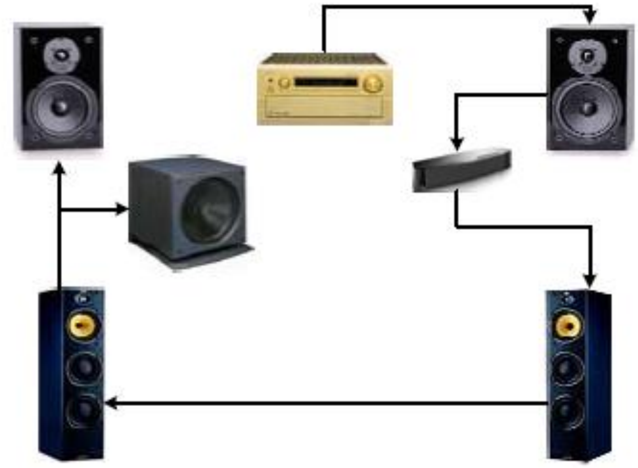


Fig b. A serial connection for a 5.1-channel system

As another application example is can include improvement in the connection ways of speakers or PA (Public Address) systems within a building, a company, or public facilities. In existing systems, a number of connection wires are required to send out audio signals to speakers from the system. Furthermore simultaneous broadcasting with multiple channels at the same time is difficult because analog audio signals are transmitted to each speaker. To solve these problems, we can build the speaker connection as shown in Figure c,. In this case, it is possible to send different messages to different speakers or to transmit some broadcasting messages to only designated speakers because digital audio signals are serially transmitted.

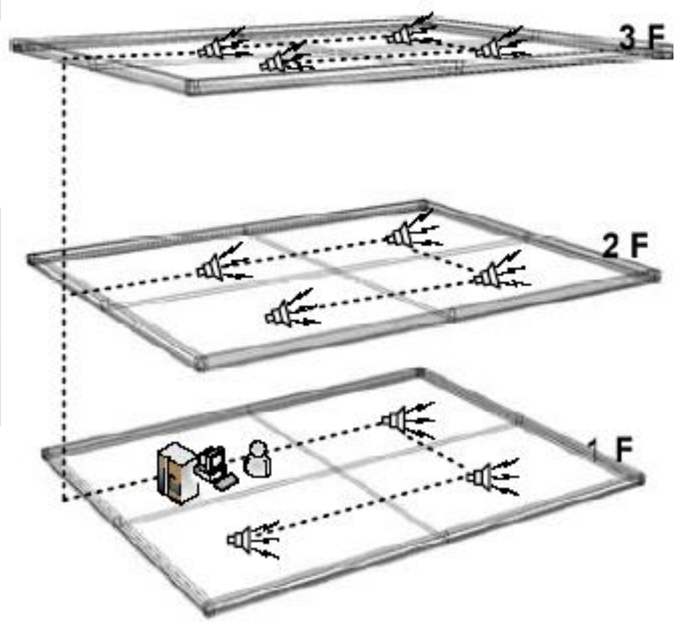


Fig c. A connection example for a PA system within a building

Audio analog signals from an audio unit need to be converted to digital signals for digital transmission. However, the conversion is not necessary if the audio unit is a digital audio device, which generates digital audio signals for the transmission, such as a HTS based on a DVD.

B. System Design

In existing audio systems with one-to-one connections between each channel and its corresponding speaker, analog signals contain driving power as well as audio data. For the high quality of a HTS, 400W power is provided to each channel through a tinned UP-OFC cable of 7 mm diameter. But converted digital signals used in the technique proposed in this paper contain audio data only. The amount of digital audio data transmitted depends on the quality of sound and the number of speakers [5].

1) Development specification

Table 1 shows the specifications to implement the serial connection system proposed in this paper. Analog signals from each channel are converted into digital signals for I^2S (Inter-IC Sound)-bus data which is a serial bus design for digital audio devices and techniques such as compact disc (CD) players, digital sound processors, and digital TV (DTV) sound[8]. For the conversion to I^2S-bus data, the PCM (Pulse Code Modulation) technique with a speed of 44.1 KHz in terms of Nyquist sampling theory is used for audio frequency 16 Hz - 20 KHz. In this case, each channel has a 24-bit sampling size. In case of digital signal outputs like DD/DTS or S/PIDF, this conversion process is not necessary [4, 8]. In the case of a general HTS with 6 speakers, each speaker requires data processing at the speed of 0.9 Mbit/sec to get as a sound quality as a CD (Compact Disk). Therefore, a 5.1- channel audio system requires at least 5.4 Mbit/sec without digital data compression. A system with a serial transmission rate of, at maximum, 9Mbit/sec for the future expansion.

TABLE 1. DEVELOPMENT SPECIFICATIONS

Functions	Spec.	Remarks
Sampling method	PCM	I^2S
Sampling rate	44.1 Khz	CD quality
Sampling data size	24 bits	Each channel
Number of speakers	6	5.1 channel
Max. TX speed	9 Mbit/sec	
Transmission method	Differential signaling	

2) Overall system design

The block diagram of the connection system Figure d. The left side of the figure converts analog audio data to digital audio data with ADCs (Analog-to-Digital Converters) and synthesizes the converted signals into packets to be sent to speakers serially. In the right of the figure, transmitted signals are identified by signal selection parts and converted into analog signals by DACs (Digital-to-Analog Converters).

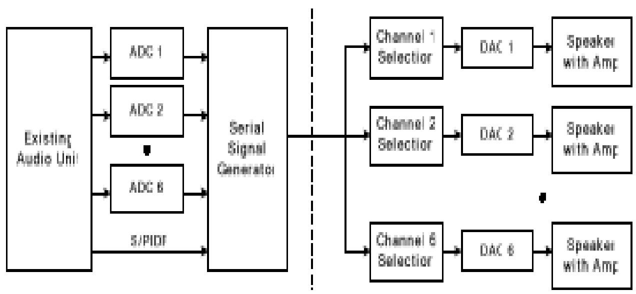


Fig d. Block diagram of the connection system

Figure e shows the visual diagram of data transmission for Figure d.

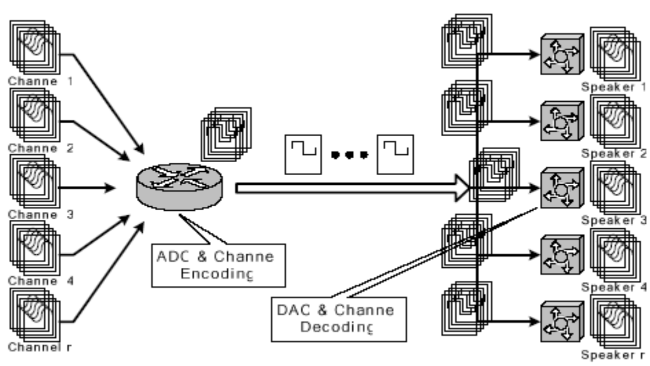


Fig e. Visual diagram of data transformation

3) I^2S-bus data format and packet generation

How to generate packets from I^2S-bus data converted by ADCs is described in Figure f. Analog signals from each channel of an existing audio unit are converted into digital signal in I^2S-bus formats (as shown in Figure f.1). Then the converted digital signals of each channel are used to generate packets (as shown in Figure f.3) by means of an encoding process.

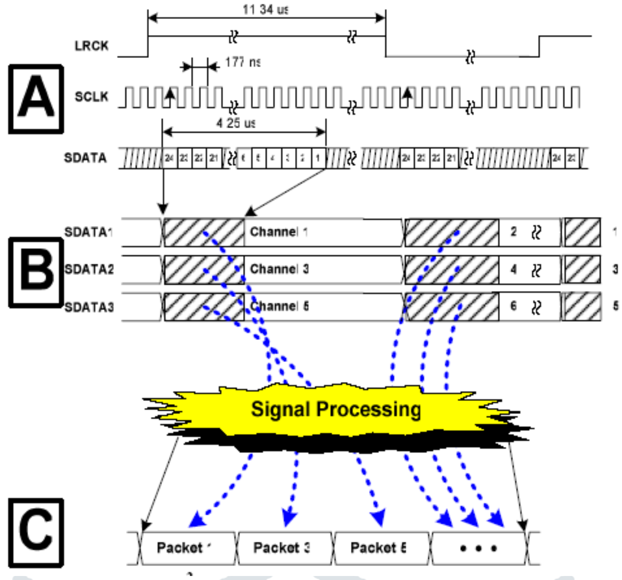


Fig f. I^2S-bus format and packet generation

As shown in Figure f, the output of 24-bit effective data begins at the second rising edge of the SCLK signal after the state of the LRCK is changed. According to the value of LRCK (either '0' or '1'), different channels are activated. That is, if LRCK is '0', then channels 2, 4, and 6 generate I^2S-bus signals at the same time, otherwise channels 1, 3, and 5 do the same task. Packets for three channels are generated every 11.34*us*. Therefore, each channel generates its packet every 22.68*us*. In this case, data processing and transmission for three channels should be done for time duration 7.09*us* (=11.34*us* – 4.25*us*), during which there is no effective data on SDATA. Relatively low bandwidth is used for long distance transmission lines. We need to compress data to send audio signals with low bandwidth lines.

4) Serial audio signal processing

The fast processing hardware technique is required to deal with data from three channels within at least 7.09*us*. I^2S-bus data for each channel are processed as shown in Figure 7. We can see the processing of sampled data at times t, t+1, and t+2.
Data sampling at a high speed is required to get audio signals with high quality. Since there is no big difference between two sequentially sampled data, if we apply an *xor* operation to the two data and scan the operation result from MSB to LSB, fast processing is possible because only a small amount of data storage is required. In the example of t+1 *xor* t2 shown in Figure f, g bits are used to process 24-bit data [6].

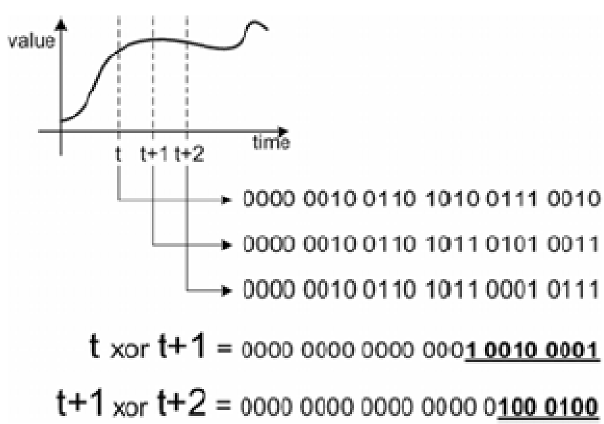


Fig g. Serial audio signal processing

5) Packet formats

The kind and format of packets used for 24-bit data of each channel is shown in Figure h. In the figure, the I point means the arithmetic average of the maximum and minimum values of voltage range for analog audio signals. Therefore, converted digital data of the I point with 24 bits are "0000 0000 0000 1111 1111 1111". If we define the I point as above, we can get fast data processing and reduce packet loss errors for multiple channel processing.

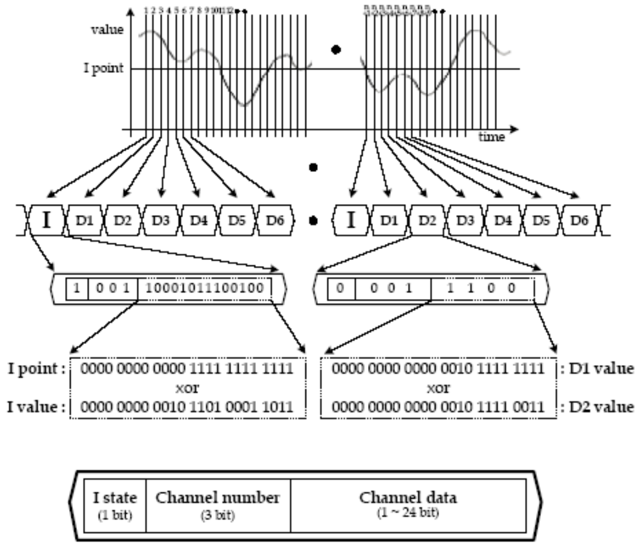


Fig h. I/D compression method and packet format

As shown in Figure h, a packet has a header part and a channel data part. The header part is divided into two parts: an I state part (1 bit) to represent either an I packet or a D packet, and a channel number part (3 bits) to indicate the number of the channel which will send data with the use of the packet. The channel data part of an I packet contains the *xor* operation value between the I point and the first sampled value. In case of a D packet, the part has the *xor* operation value between two consecutive values. If packet loss happens during the transmission of a D packet, the recovery operation begins from the following I packet. I packets are sent at regular intervals and the frequency of sending I packets can be determined according to transmission conditions. The total size of a packet has the range of 5 bits to 28 bits.

6) Design of a serial signal generation part

The detailed block diagram for hardware implementation of the above mentioned job is shown in Figure i, which corresponds to the serial signal generator block of Figure d.

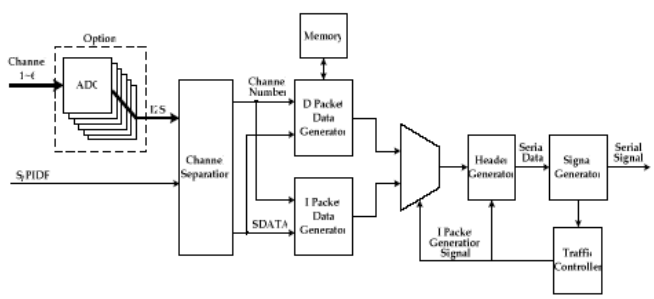


Figure i. Detailed diagram of data generation for serial transmission

Let us explain the diagram in detail. Analog signals from each channel are sent to the Channel Separation block after converting into I^2S-bus data formats with an audio ADC. In this case, digital audio signals (like S/PIDF) are sent to the block directly. The Channel Separation block separates SDATA data of each channel from mixed I^2S-bus data signals and sends them to Packet Data Generator blocks, which are divided into D and I blocks to generate D packets and I packets respectively. The Traffic Controller block produces packet generation signals to control their transmission rates according to transmission conditions. That is, according to this signal, the generation of I packets or D packets is determined.For example, consider that an I packet is generated every 20 packets in the digital audio signals with 44.1 kHz sampling speed. Signals can be recovered within, at most, 0.41us in case of a packet loss. Since a 5.1-channel audio system needs a smaller amount of data than the transmission bandwidth, and the Traffic Controller block generates many I packets, human beings can not recognize the packet loss.

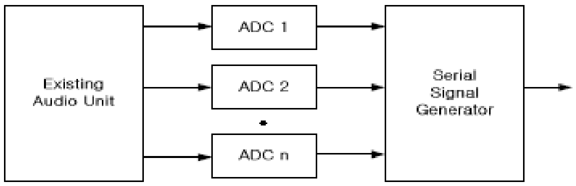
7) Differential Transmission Circuits

In general, the transmission of digital signals uses differential signal transmission circuits which take advantage of voltage difference on wires. They have good characteristics for communication distance and speed, as well as noises. Differential signal serial transmission guarantees the transmission distance of 1.2 Km and bandwidth of 10 MHz [7]. Therefore, we do not have to add an amplifier circuit during signal transmission to expand the communication distance in relation to the installation of more speakers for a HTS. In this paper, UTP (Unshielded Twisted-Pair) cables are used to send digital audio signals to serially connected speakers.

IV. MERITS AND DEMERITS OF SERIAL SPEAKER CONTROL

Compared to existing audio systems with the use of analog audio signal transmission, the proposed technique, which allows serial speaker connection, has many advantages. They include easy installation and construction, cheap connection costs with less connection wires, no distortion in audio signal transmission, and easy management and maintenance of speakers. This technique transmits and processes audio data at the speed of 5.4Mbit/sec to guarantee quality as good as a digital CD level.The disadvantage is that we need to provide each speaker with additional things like driving power, a digital-analog converter, and an amplifier . Multi-channel speaker system that connected each speaker serially in the PA (Public Address) system. Disadvantages of the system are the increased, size of system since external independent audio systems are required and the requirement for physical switching between various analog sources. Due to fixed hardware architecture it is difficult to change the number of channels. Thus, flexibility is reduced because of the limited number of channels for a compact system.

SERIAL SPEAKER WITH SOME MODIFICATIONS AND OVERCOMING ITS DISADVANTAGES WAS USED FOR SMART SPEAKER CONTROL OF HOME AUDIO NETWORK. WHICH IS EXPLAINED BELLOW.



Transmission-part block diagram of the serial system

Gradually, PCs are moving into living rooms, as most audio sources are becoming digitized and various types of existing digital audio sources are best played back using comprehensive software running on a computer. The digitization of a living room multimedia environment can be more efficiently extended to the entire home environment by replacing the multiple-wired speakers with single-wired speakers using digital control. In this paper, we propose an intelligent audio system controlling multiple speakers through a network cable and a USB cable communicating with the living room PC as shown in Fig. 1. The PC is connected to the proposed main controller using the USB with less than 5m in distance. The main controller communicates with each of the

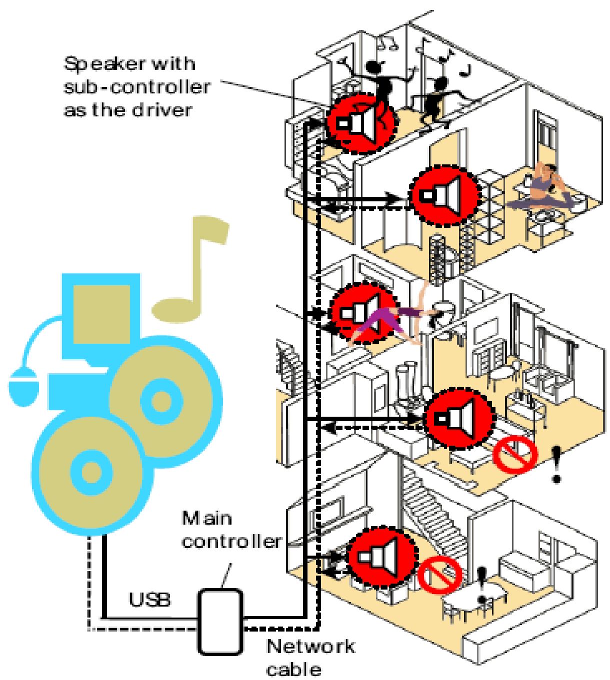


Fig. 1. Proposed setup of an intelligent home network audio speaker system.

sub-controllers using a serial network cable. The sub-controllers either drive the speakers or monitor the rooms. The control of digital data streams through a centralized PC allows easy management of multiple sound sources at a much lower cost. In addition, a channel can be added simply by linking a sub-controller that controls a speaker with a different channel ID in a daisy chain fashion, instead of rewiring additional line from the analog amplifier to the speaker and upgrading the expensive central analog amplifier channel. The additional channel is activated by programming the main controller to issue audio packets for the particular sub-controller. We have previously developed a serial multi-channel speaker system [1-2] without the USB interface and demonstrated its advantages over analog systems. However, without the USB interface, providing multiple audio sources through a single console was difficult to achieve. In addition, in order to enable bidirectional communication capability, a well-established protocol was necessary. We chose USB as the communication channel between a PC and the proposed main controller due to its high bandwidth and robust interface. USB is the standard which is widely used. In particular, USB 2.0 has high transmitting speed, and it is generally used for portable storage devices and multimedia consumer products. USB has four transmission methods: control, bulk, interrupt and isochronous transmission. Isochronous transmission is suitable for real time transmission because during this mode, error correction operations do not occur and thus deterministic transmission rate can be achieved. Even if the error correction is disabled during isochronous operations, the near proximity of less than5m between the PC and the main controller ensures the errors are negligible. Using USB 2.0 high speed isochronous transfer, supporting max speed up to 24.5765Mbytes/sec On the other hand, bulk transmission method is used for channel selection data transfers because this does not require real-time operation while errors can be critical. In the case of isochronous transfer, even if an error is detected after data is sent, the host does not re-transfer. Because of this, handshake packets are omitted. In the case of audio or video that is played in a real time, because it is very difficult to discern one or two bit error using person's ears or eyes isochronous transfer is the best choice. In the case of bulk transmission, it is good for data transmissions where timing is not important. Because it waits until other transfers are done, it sends data without bus interrupt. Thus, the system transfers audio data by isochronous transfer and control data for channel selection by bulk transfer.

The main advantage of utilizing digital technology to the speaker system is the bidirectional communication capability between the main controller and the sub-controllers. For example, a request from the main controller to a designated sub-controller for information can be issued and corresponding result can be received in a handshaking manner. In this paper, this type of communication is explored for the serial speaker system to intelligently provide suitable music for different rooms. An extension to this proposed technique also allows multiple tracking of the movements of persons to provide continuous stream of music catered to individual needs. Fig. 2 shows a simple hardware configuration that would allow an intelligent home network audio system.

The sub-controller through its digital input port can receive a motion detector result and inform the PC whether the room is occupied or not. A more sophisticated interface would use

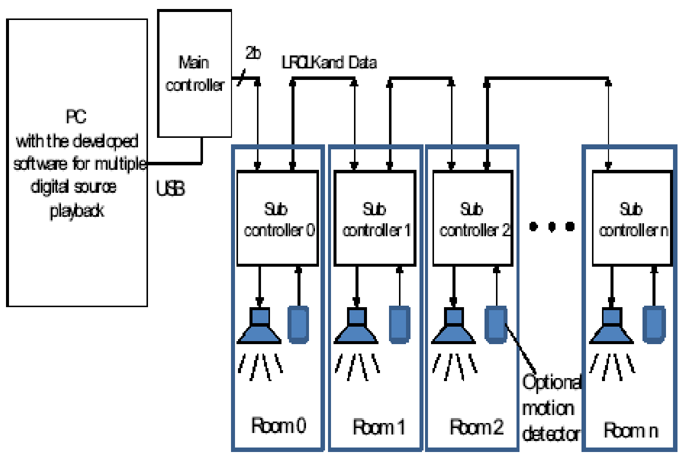
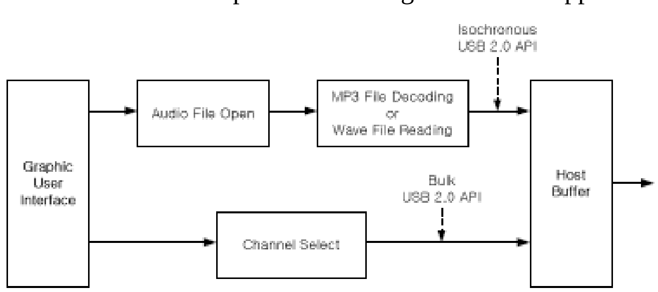


Fig. 2. A simple configuration for intelligent home audio network system

RFID readers in each room and let them figure out exactly who has entered the room, provided that the individual is tagged. Nonetheless, it is up to the software in the PC to decide who is in the room and which audio stream to be played. In this paper, we focus on the hardware that allows the proposed smart speaker control possible. Moreover, the Fig. 2 only shows a single channel per room for demonstrative purposes. Each room can have more than a single channel provided that the room has more than one sub-controller speaker system.

IV. Architecture

The architecture of the proposed hardware framework can be categorized into two parts. The first part is the USB interface between PC and the main controller and the second part is the serial network interface between the main controller and the sub-controllers. Fig.3 shows the detailed interface for the first part. And the digital sources supported are wave files and mp3 files.



PC application block diagram

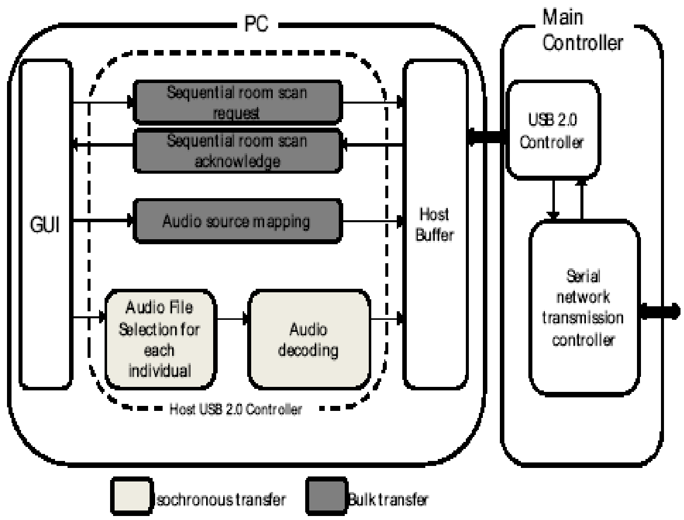


Fig. 3. Interface between the PC and the main controller

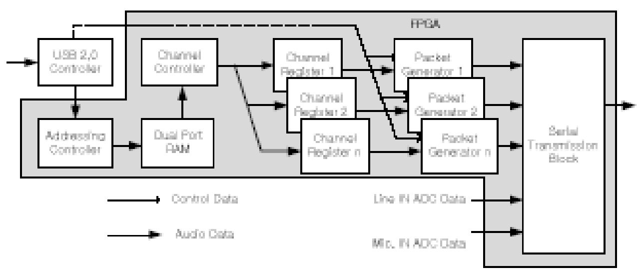


Fig m. Internal block diagram of FPGA (main controller)

The FPGA operates with data received from the USB controller as shown in Figure m. The flow of audio signals is represented by a solid line, and flow of control signal for channel selection is represented by a dotted line. An addressing controller defines addresses to save sampling data received from USB controller to a dual port RAM. And Channel controller reads saved data from dual port RAM for a sampling interval. Data stream must not be stalled for continuous playback audio signals. Thus, the RAM should be updated and be read at a real time. Data is transmitted to speakers serially after compression and to be packetized by packet generator after data read from RAM is saved to each channel register. The system can transmit control data by using bulk transfer to select the chosen channel. The control data to selected channel set up packet's ID.

System Design Specification

Table 1 shows the specifications of the serial connection speaker system proposed in this paper. Using USB 2.0 high speed isochronous transfer, because of supporting max speed up to 24.5765Mbytes/sec, it is possible to send multi-channel audio signals in a real-time fashion. In addition, we configured the system to select a channel using bulk transfer. DAC in the receiving part, sampling frequency is set at 48 kHz with each channel defined with 24 bits sampling size. And the digital sources supported are wave files and mp3 files.

Table 1. Development specifications

Function	Specification	Remarks
USB interface	2.0	Supported high speed
USB transmission methods	Isochronous	Audio source data
	Bulk	Channel selection
Sampling rate	48 kHz	CD quality
Sampling data size	24 bits	Each channel
Supported files	Wave, MP3	

The speakers are activated based on periodic sequential room scan operations initiated by the PC. Each of the rooms acknowledges with a flag bit indicating whether the room is occupied or not. Based on the occupancy, the application can be programmed to play appropriate audio streams by mapping the preloaded audio sources with the occupied room.

The PC processes mp3 files or wave files preloaded audio sources. A PC application writes data to USB transfer buffer using API functions after decoding the files. The audio data is decoded to 24-bit pulse code modulation (PCM) format. In the transmission part, PC transmits many sound source data to the USB controller and it transmits channel data to main controller according to the protocol.

The sampling frequency of the audio data is 48 kHz, therefore the PC must send 48,000 x 3 bytes per second for each room. In the case of isochronous transfer in USB 2.0, the host can transfers packets up to 1K bytes at a time with an interval of 0.125ms. During one second, the host can send 8000 packets each containing 6 sampled data in a packet. Figure L shows process of isochronous packet transfer in PC applications. The first packet carries 1st-6th sampling data, and the second packet carries 7th-12th sampling data. With this process, 47,995th-48,000th sampling data are sent at the last. This process continues for one second, and is repeated continuously. Consequently the maximum number of rooms supported is 56. However, the actual supported number of rooms is less and is more critically dependent upon the number of channels per room. For instance, if there are stereo channels for each room, the maximum number of supported rooms will drop down to 28. In order to provide sustained rate of 48 kHz PCM, dual port memory is required within the main controller. This ensures simplified real-time processing of audio data at each of the sub-controllers.

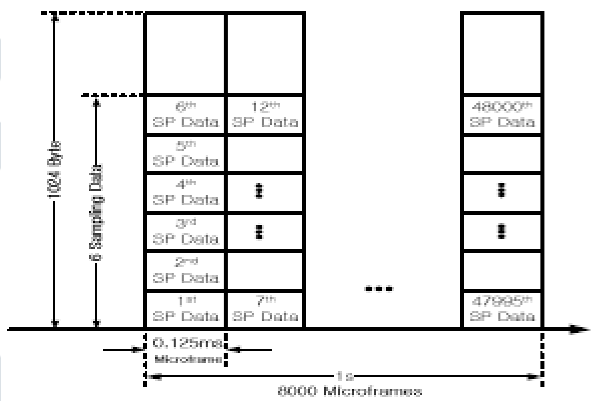


Fig L. Isochronous packet transmission process

Need for dual port memory

The most important criterion for the system is to play audio signals in a real time. So, channel data received from PC is sent to speaker at 48 kHz sampling frequency without omission. Because 1,024 bytes of data are sent for every 0.125ms, buffers or memories are required to meet the rate of 48 kHz. Because the data is updated and should be read in real-time, dual port RAM is used to continuously support burst writes and 48 kHz reads. The dual RAM interface is shown in Figure j. An addressing controller chooses address to save data that is received USB controller from 0x0000h on the dual port RAM. And then, data should be read from 0x0000h on a dual port RAM at a sampling rate. It takes one second to read data from 0x0000h to 0xBB7F on the dual port RAM, and this process is repeated continuously.

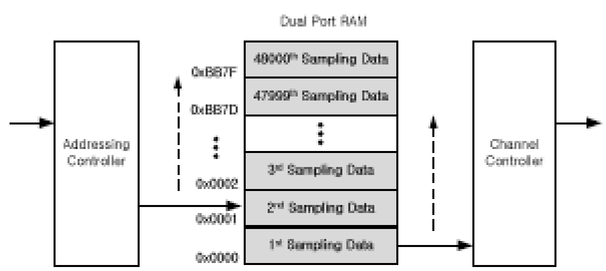


Fig j. Dual port RAM interface

Bandwidth of the network cable side can be maximized using a compression scheme. The compression scheme used is delta coding. The delta codes of each channel were combined with a fixed rate of PCM data to limit the error accumulation. The intermittent PCM (29 bits, including the header) data corrects any noise that affected the channel during delta coding or data or data overflow which depends on the number of bits used for delta coding (4 bits for the difference; 9 bits including the header for this implementation). Fig. 4 shows the detailed description of the serial network transmission controller within the main controller. The PCM data rate can be dynamically adjusted for optimal rate to strike a balance between data compression and audio quality. In this implementation, however, the PCM data rate was fixed to every 20 packets per channel. Therefore, an average of 10 bits per sample is achieved.

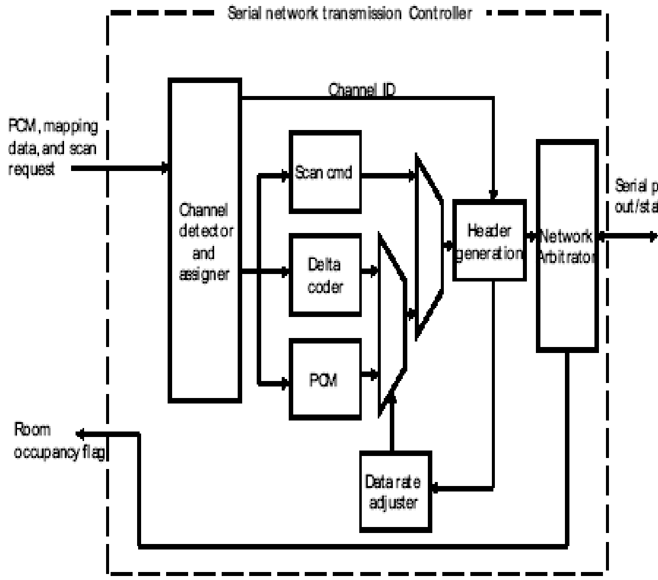


Fig. 4. Serial network transmission controller

The second part of the framework involves interactions between the main controller and the sub-controllers. This part of the hardware interface is similar to the result we have presented for serial channel speaker system. Before transmitting data serially, sampling data read from dual port RAM is compressed and packetized. Figure f illustrates the process. Data read from RAM is saved in the channel register. And that data is compressed in compression block of packet generator. The channel data and that is sent using bulk transmission sets packet's ID and a packet is generated with compressed data and the ID. The serial transfer block transmits a packet serially for each channel. Figure n shows a flow of serial transmission.

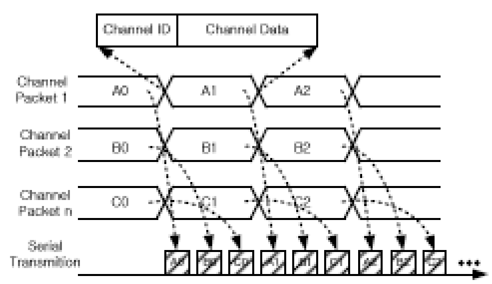


Figure n. Packet transmission process

Each channel packet is generated at the rate of 48 kHz and is sent serially and sequentially before a next packet is generated. Because all speakers are connected through a single line, each speaker verifies all packets' IDs, and chooses and reads a packet that has the corresponding speaker's ID. Then, the speaker decompresses the compressed data and the DAC is used to reproduce the audio signal. Packet regeneration process is shown in Figure p.

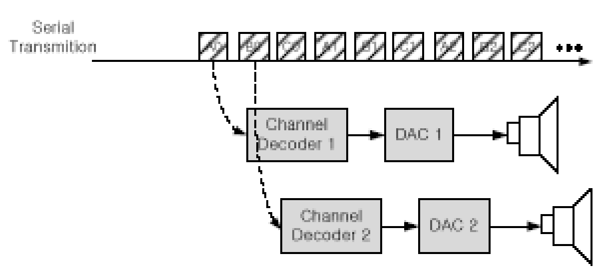


Figure p. Packet regeneration process

The communication protocol does not use error correction capability since it is operated at a relatively low frequency. The only difference is the digital input provided to the subcontroller for room monitoring function. When a scan is requested by the PC, each sub-controller receives the scan command. As soon as the scan command is sent, the main controller switches to a receive mode and waits for three clock cycles before returning to a transmit mode. The scan command is detected at the sub-controller when the 2-bit of MSBs for the 5-bit header value is ‘00’. The following 3 bits are used to select the sub-controller to be scanned. When the header MSBs are ‘01’ the data is PCM and if it is ‘10’ the data is delta code. Upon detection of the scan command, only the selected sub-controller switches to the transmit mode and sends the digital bit that was received at the digital input of the sub-controller. The unselected sub-controllers regard it as NOP. Three clock cycles after transmitting the value, the subcontroller returns to the receive mode. Redundant clock cycles were used to ensure proper latching of the digital bit at the receiving end of the main controller. Fig. 5 shows the timing diagram showing the bidirectional communication between the main controller and the sub-controller for an 8-channel configuration.

Note that for both the main controller and the sub-controller,PLL in the FPGA is configured to generate 4.8MHz clock from LRCK. So, in order to maintain 48 kHz rate, the data

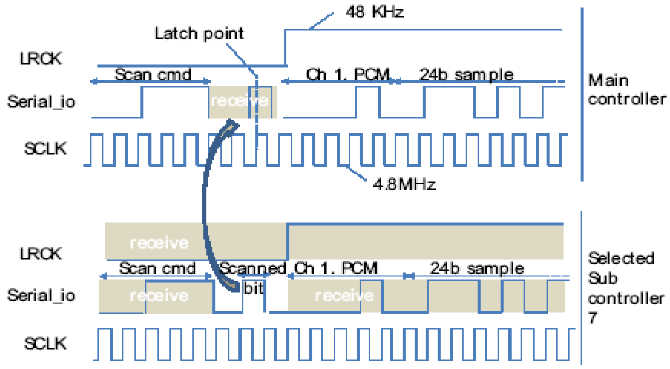


Fig. 5. Bidirectional communication while satisfying 48kHz sample rate.

packet must be less than 100 bits. We designed the compressor such that there are a maximum of one intermittent PCM data per LRCK cycle. This means that maximum bits transmitted per LRCK cycle is 92 (= 29x1+9x7) bits for sending different streams to 8 channels. Thus, there are 8 SCLK cycles to spare at the minimum before the next set of sampled data arrives. We used the 8 SCLK cycles to receive and transmit the scan operations such that this does not interfere with the 48 kHz audio transmissions. If more channels are introduced, the SCLK frequency must be configured to be higher, and the need for incorporating error correction capability may arise.

III. Results

To demonstrate the intelligent home network audio system, we constructed a simplified configuration where 4 subcontrollers are connected to the main controller. As shown in Fig. 6, the main controller is composed of a processor core with an embedded USB 2.0 controller and an FPGA chip. The processor core performs the enumeration process. For I/O, we constructed pipes with an isochronous OUT endpoint and a

Fig. 6. The main controller board.

Fig. 7. The sub-controller board.

bulk IN/OUT endpoints. Fig. 7 shows the sub-controller board. The board has an output interface for the speaker. The speaker is driven by the DAC output which converts the decompressed PCM data from the FPGA into analog audio signal. In addition, the digital input is provided directly to an FPGA input pin such that an external motion detector can interface to it. The actual test setup involved 4 sub-controllers driven by a main controller connected through USB with a notebook computer. For the demonstration, the motion detector output is emulated by an FPGA board. The FPGA board outputs 4 digital signals and each of the signals are connected to the digital inputs of the 4 sub-controllers. The notebook scans the data from each of the sub-controllers periodically and updates the occupancy information. If the room is determined to be occupant, then an audio source

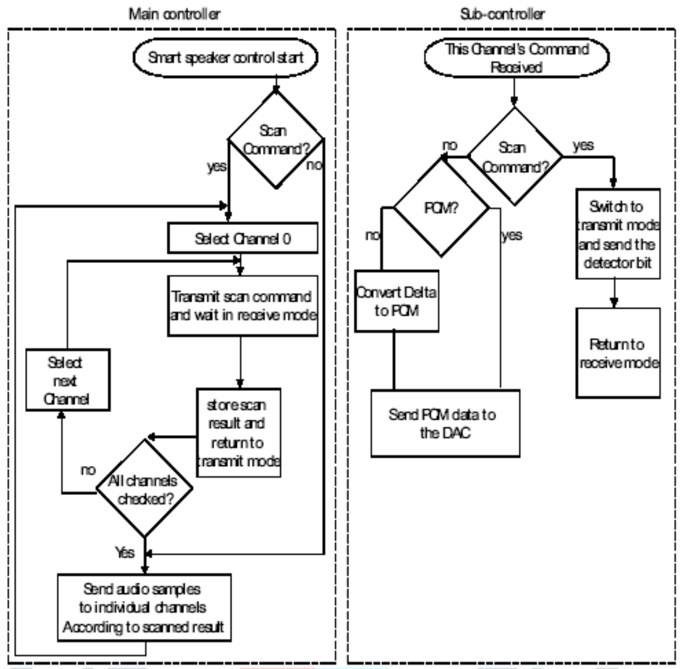


Fig. 8. Flow chart showing the operation of the controllers implemented for the demonstration.

assigned to the room is played. The operation of the main controller and sub-controllers are illustrated in Fig. 8. The main controller, which is directly controlled through the USB bus, receives the data frames from the PC. The USB controller within the main controller board and the FPGA are used to extract the channel information, scan commands, and the raw audio data. The bulk transfer mode is used to both send and receive the scan commands and the scan results respectively. As shown in the flowchart, when the main controller receives a frame data, it first checks for a scan command. If there are no scan commands, the audio sources are sent using the channel data and the previously stored occupancy information for each room. Upon detection of the scan command, all channels are sequentially requested for the scan result by issuing a scan command to the sub-controller. Note that individual channel scan information is requested for an audio sample period. Therefore to scan all possible rooms, at least four data sample periods of audio signals are spent, provided that there were no errors during the bulk transfer. During the scanning operation both the main controller and the sub-controller that received the scan command switch between receive and transmit mode within 625ns. Since the length of the cables between the subcontrollers are short in this demonstration, we may actually require slightly higher SCLK rate to provide more spare cycles for the bidirectional transitions. The selected sub-controller simply decodes the headers to determine whether the packet is a PCM data, delta code, or a scan command. If the header indicates that it is a scan command, the data from the external input of the sub-controller is sent back to the main controller. If it is a delta code, it is decompressed back to PCM and it is sent to the DAC. Although scanning operations need not be frequent, we provided scanning command on each of the USB frames, or every 0.125ms. A push button on the FPGA board emulating the motion detectors was used to verify that correct room was identified. Also in this demonstration, we assumed it is a single floor house. The GUI in Fig. 9 indicates that room 3 representing the third sub-controller from the top is provided

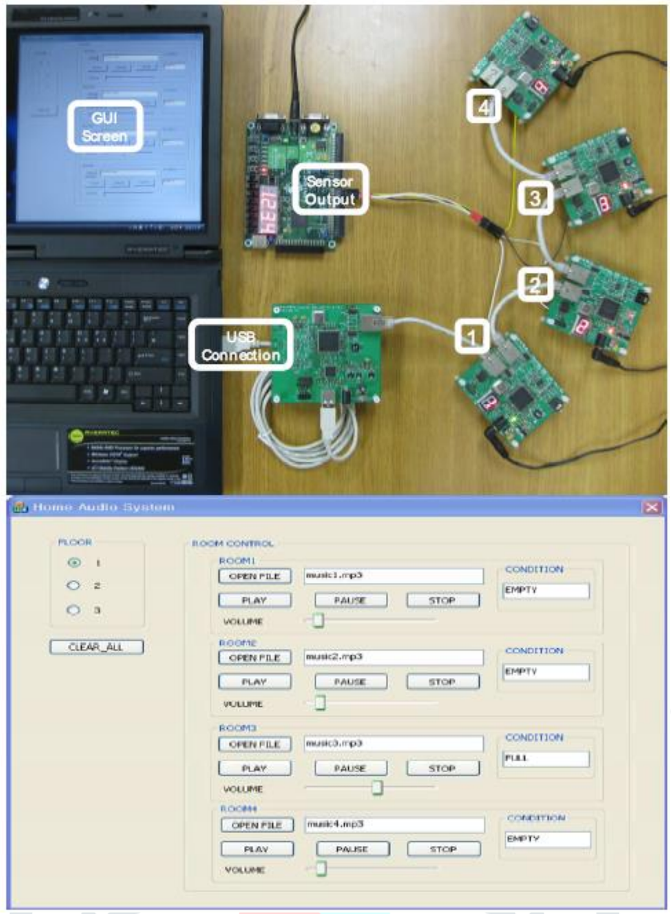


Fig. 9. Demonstration of smart speaker control for home audio network.

with digital motion detection signal 'high' and the rest of the sub-controllers are provided with 'low'. Although more sophisticated algorithm for audio source selection algorithms could have been developed, our focus was on the hardware framework that makes the intelligent home audio possible. the word it modifies, usually without a hyphen. There is no period after the "et" in the Latin abbreviation "et al." The abbreviation "i.e." means "that is," and the abbreviation "e.g." means "for example." An excellent style manual for science writers is [7].

IV. Conclusion

Speaker control system for a home audio network is demonstrated using a bidirectional communication capability.The proposed USB connection between the main controller and a home entertainment PC allows an intelligent audio control for individual rooms in the house. In addition, we proposed a simple protocol to provide bidirectional communication capability between the main controller and the sub-controllers without interrupting the continuous 48 kHz sampling rate for each of the rooms with varying numbers of speaker channels. The PCs are becoming a living room appliance that we rely upon for various home network functionalities. In such an environment, the proposed audio system is a cost effective and value added approach compared to the analog speaker connection system being used today.

References

[1] Moovin Song, Doo Young Yi, Ohkyun Kwon, Yunmo Chung, "Serial Connection of Speakers for Multi-channel Audio System," Proceedings of ISEE 2003, pp.440 – 443, Japan, November 2003.

[2] Moovin Song, Ohkyun Kwon, Yunmo Chung, "A Serial Connection Technique of Speakers for Multi-channel Audio systems," IEEE Transactions on Consumer Electronics. Vol. 51, No.2, pp. 611-616, May 2005.

[3] Sunyong Lee, Taewan Kim, Jinsung Byun, Yunmo Chung, and Moonvin Song, "Intergrated Control System of Multi-Channel Audio Signals for Serial Transmission," IEEE Asia Pacific Conference on Circuits and Systems, pp.1446-1449, December 2008.

[4] C. Busbridgem Y. Huang, and P. A. Fryer, "Crossover Systems in Digital Loudspeakers," AES Journal, Vol. 50. No. 10, pp. 791, 2002.

[5] B. Forouzan, Data Communications and Networking Third Edition, McGraw Hill, 2003.

[6] W. Wolf, Modern VLSI Design Third Edition, Prentice Hall PTR, 2002.

[7] http://para.maximic.com/compare.asp?Fam=RS485&Tree=Interface&HP=Interface.cfm&ln.