Recent Progress in Digital Audio Technology

TOSHI T. DOI

Sony Corporation, Tokyo, Japan

This paper reports on recent progress in digital audio technology, including improvements in analog-to-digital (A/D) and digital-to-analog (D/A) converters, magnetic-recording-head design, modulation schemes, error-correction schemes, and editing processes. The improvements are remarkable and have contributed greatly to cost reduction, better sound quality, reliable higher packing density, and increased versatility. Newly developed digital audio systems that include professional recorders, satellite broadcasting, and digital audio disks are described. Large-scale integrated (LSI) circuits developed for consumer systems and their impact on digital audio are also reviewed.

1 BASIC TECHNOLOGY

1.1 A/D and D/A Converters

Most converters used heretofore depend on the accuracy of resistor networks. The resistors representing the most significant bit (MSB) or the second significant bit are critical. Variations due to heat or shock spoil the monotonicity of conversion, thus reducing sound quality considerably.

Several converter designs guarantee monotonicity. One is the integration converter, which is not considered applicable at sufficiently high speed for digital audio.

Fig. 1 shows monolithic A/D (CX-899 [1]) and D/A (CX-890 [2]) converters, which obtain 16-bit accuracy without trimming. They are based on dual-slope integration, which does not always guarantee monotonicity because the balance between currents I_0 and i_0 shown in Figs. 2 and 3 must be maintained precisely. The required precision is easily achieved without adjustment. Table 1 shows the specifications for both converters.

1.2 Head Technology

Fig. 4 shows a magnetic head for professional recorders, which contains 26 tracks over a 0.5-in (12.7-mm) width. The material is single-crystal ferrite. The structure of this head is identical with that of a video

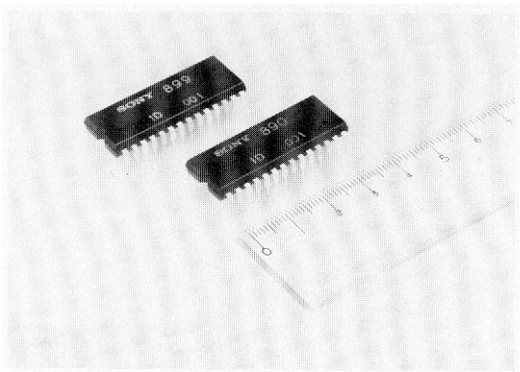


Fig. 1 Examples of monolithic A/D (above) and D/A (below) converters.

head, and it is compared with a conventional audio head in Fig. 5. Side gaps are eliminated, increasing efficiency and lowering crosstalk. Shielding between tracks is not necessary, thus simplifying the structure.

Fig. 6 shows the crosstalk versus guard band for different core widths.

Fig. 7 illustrates another new head technology, namely, thin-film heads developed for a compact-cassette digital audio deck [3]. The upper head is a multiturn head for recording, the lower an MR-type head for playback. Thin-film heads have the following general features:

1) The guard band of the recording head can be reduced greatly because crosstalk is extremely low and the wiring space is not necessary.

2) The bandwidth of the playback head can be reduced because, at tape speeds less than 2 m/s, the signal-to-noise ratio of the MR-type head is better than that for conventional heads.

3) The head is very precise because it is produced by photolithography.

Consequently thin-film heads provide the opportunity to increase track density greatly while maintaining the same level of tape interchangeability.

1.3 Modulation Technology

Table 2 and Fig. 8 show the parameters and waveforms of conventional channel codes (modulation schemes) developed basically for computer peripherals. Early digital audio machines used one of these codes, but after a comprehensive study several new codes were developed and contributed to improved packing density [4]. Their parameters and waveforms are shown in Table 3 and Fig. 9. An example of the coding rule HDM-1, high-density modulation, is shown in Fig. 10.

1.4 Error-Correction Technology

Fig. 11 shows an example of error distribution in an optical disk system. After fingerprints had been marked on the entire surface of the disk, the error rate doubled and burst lengths increased. The same results, but more

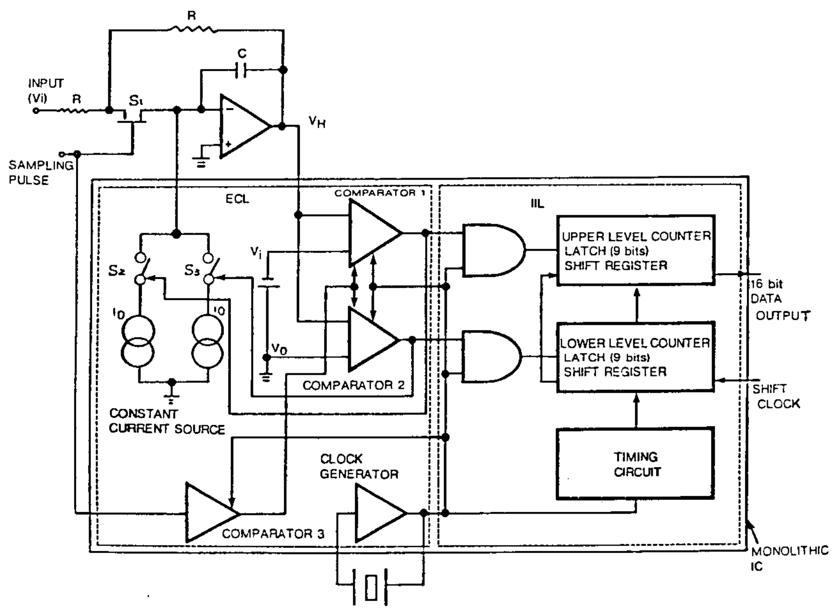


Fig. 2. Block diagram of A/D converter CX-899.

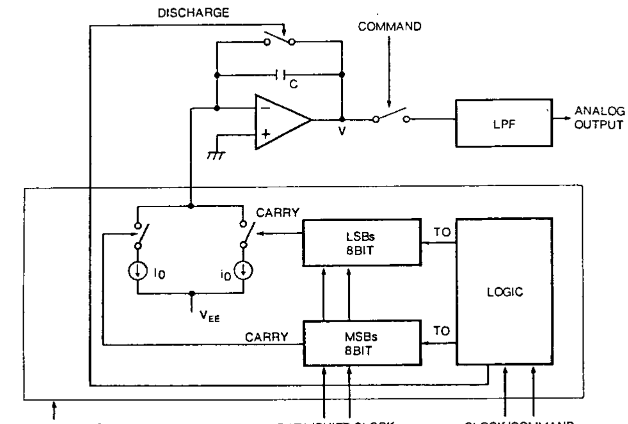


Fig. 3. Block diagram of D/A converter CX-890.

Table 1. Basic specifications of monolithic converters.

Item	A/D Converter CX-899 [1]	D/A Converter CX-890 [2]
Quantization	16-bit linear	16-bit linear
Sampling rate	Up to 50 kHz	Up to 50 kHz
Clock rate	50 MHz	33 MHz
Distortion	0.0028%	0.0022%
Power dissipation	450 mW	430 mW
Supply voltage	5 V	5 V
Process	I^2L/ECL	I^2L/ECL
Number of elements	990	1020
Chip size	2.1 × 2.3 mm	1.9 × 2.4 mm
Number of pins	28	28

severe, were obtained with a disk that had been scratched by rubbing it on a wooden table. The phenomenon is evaluated by plotting the error rate against error correlation. Fig. 12 shows such a plot for the digital audio stationary head (DASH). (See Section 3.)

A new error-correcting method named *cross interleave* was developed by the author especially for digital audio application [5]. It features high correctability, low redundancy, and less hardware. The basic principle, shown in Fig. 13, is essentially two independent block codings shuffled by a delayed interleave.

When both block codes are single-erasure correctable, it is called *cross interleave code*, the correctability of which is triple erasure. The code with optimized shuffling is called *improved cross interleave code*, which is equivalent to fifth-erasure correction, while the redundancy is equivalent to double-erasure correction of the block code (Reed–Solomon code or *b*-adjacent code).

For the cross interleave Reed–Solomon code (CIRC) adopted for the Compact Disc (CD) system, both codes are Reed–Solomon.

Fig. 14 shows random-error correctability of various codes. The curve for CIRC (by simple decoder) is not directly comparable to the others because its horizontal axis is the symbol error rate (symbol = 8 bits) and its vertical axis is the word error rate (word = 16 bits). Both axes represent the block error rates (block = 288 bits) for the other codes. Table 4 shows burst-error protection for some of the well-established formats.

1.5 Editing Technology

Electronic editing has been introduced to digital audio and is already very popular. Fig. 15 shows one of the electronic editors for a two-channel system [6]. A large memory is built in for the following purposes:

1) Data around the editing point are memorized in

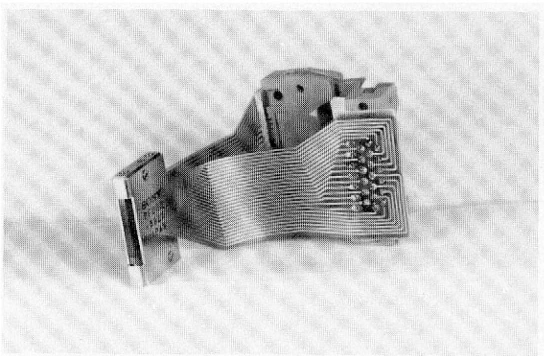


Fig. 4. Single-crystal ferrite head for professional recorder; (26 tracks in 0.5 in (12.7 mm).

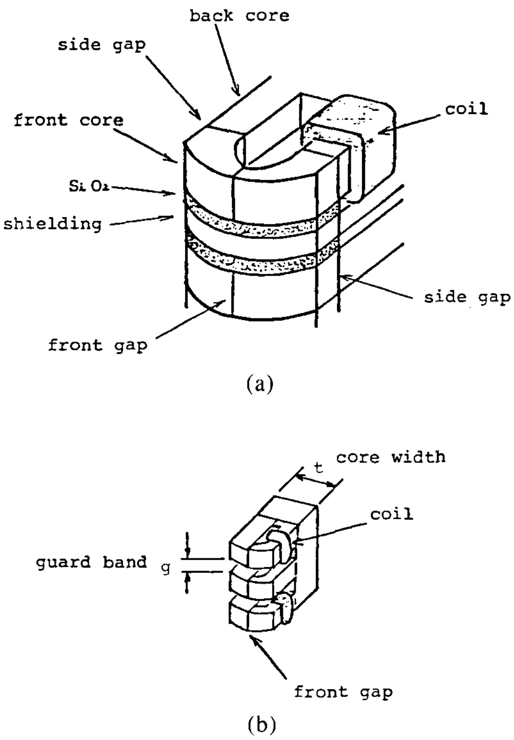


Fig. 5. Head structure. (a) Conventional head structure. (b) Newly developed single-crystal ferrite head.

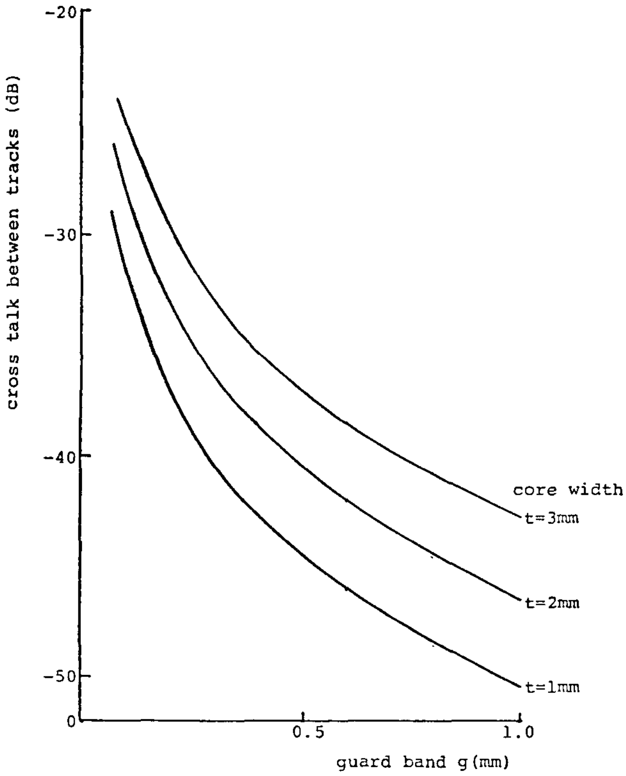


Fig. 6. Crosstalk.

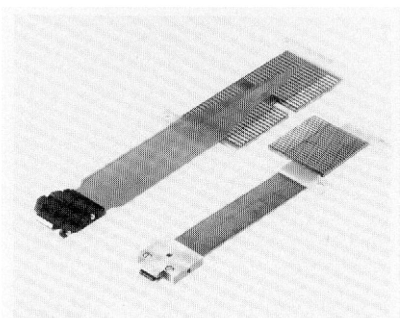


Fig. 7. Multiturn thin-film head (above) and MR-type thin-film head (below).

Table 2. Parameters of conventional channel codes.

Item	Symbol	NRZ NRZI	PE FM	MFM	Miller^2	ZM	4/5 MNRZI	3 PM
Window margin	T_w	dc	T	$0.5T$	$0.5T$	$\simeq 0.5T$	$0.8T$	$0.5T$
Minimum transition	T_{min}		T	$0.5T$	T	$\simeq T$	$0.8T$	$1.5T$
Maximum transition	T_{max}			T	$2T$	$\simeq 3T$	$2.4T$	$6T$
DC content			Bad	Good	Bad	Good	Bad	Bad
Constraint length	L_c	T	T	T	$3T$	Depends on redundancy	$4T$	$9T$
Clock rate	CLK	No	$2/T$	$2/T$	$2/T$	$2/T$	$1.25T$	$2/T$
Density ratio	DR	1	0.5	1	1	$\cong 1$	0.8	1.5
Figure of merit	$T_w \times T_{min}(T^2)$	1	0.25	0.5	0.5	0.5	0.64	0.75
Max–min ratio	T_{max}/T_{min}		2	2	3	2	3	4

T—Length of 1 bit of original.

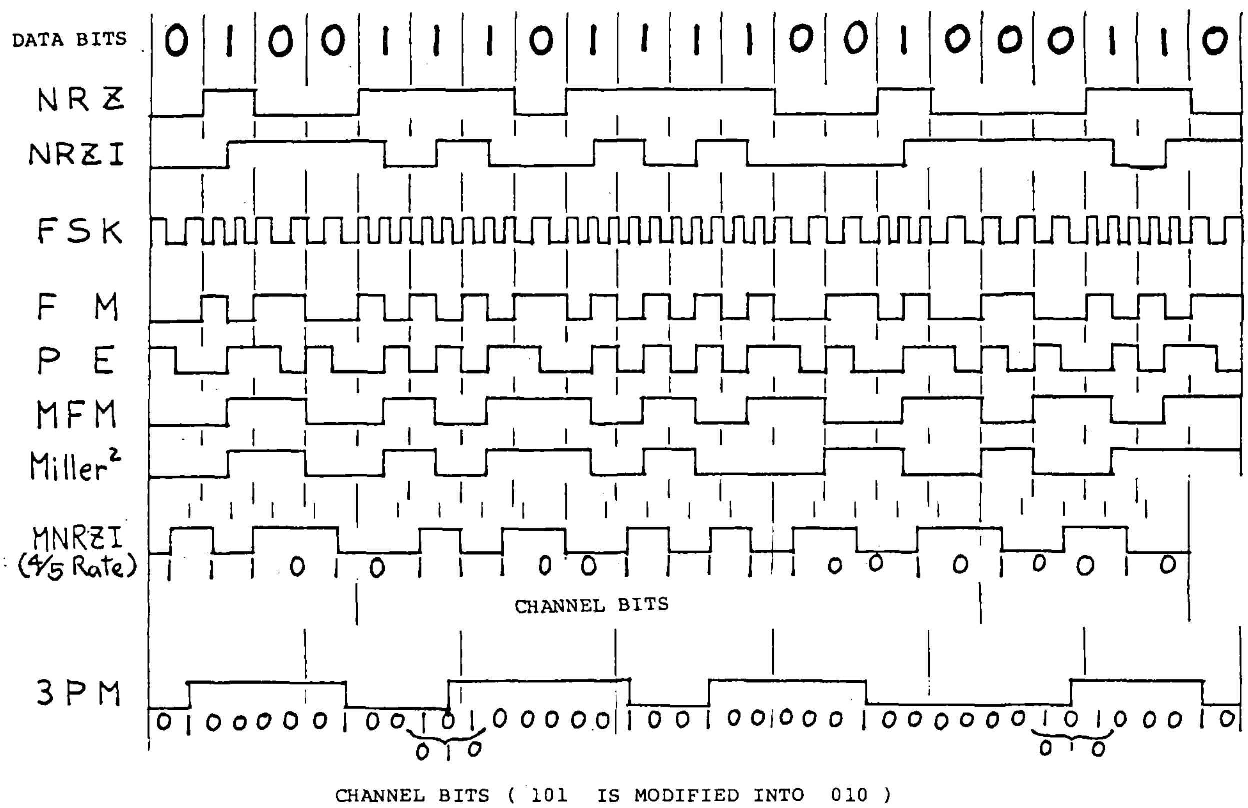


Fig. 8. Waveforms of conventional channel codes.

Table 3. Parameters of newly developed channel codes.

Items	Symbol	HDM-1	HDM-2	HDM-3	HDM-0	EFM
Window margin	T_w	$0.5T$	$0.5T$	$0.33T$	$0.33–0.5T$	$0.471T$
Minimum transition	T_{min}	$1.5T$	$1.5T$	$2T$	$1.5–2T$	$1.41T$
Maximum transition	T_{max}	$4.5T$	$4T$	$8.33T$	$4–8.33T$	$5.18T$
DC content	dc	Bad	Bad	Bad	Good	Good
Constraint length	L_c	$5.5T$	$7.5T$	$12T$	$3–12T$	$8T$
Clock rate	CLK	$2/T$	$2/T$	$3/T$	$2/T–3/T$	$17/8T$
Density ratio	DR	1.5	1.5	2	1.5–2	1.41
Figure of merit	$T_w \times T_{min}(T^2)$	0.75	0.75	0.67	0.67–0.75	0.664
Max–min ratio	T_{max}/T_{min}	3	2.67	4.17	2.67–4.17	3
Remarks		From [4]		$m=4$, $n=12$, nonlinear transformation	DC content is improved by adding some redundancy	From [7]

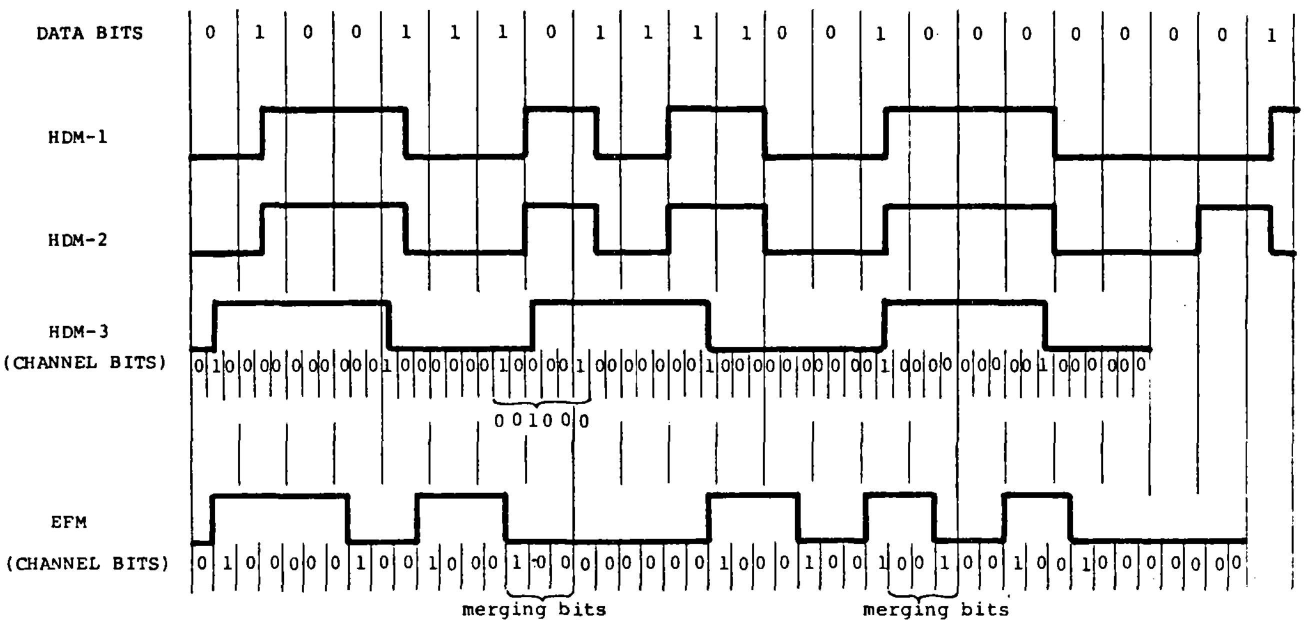
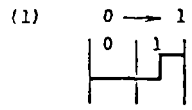
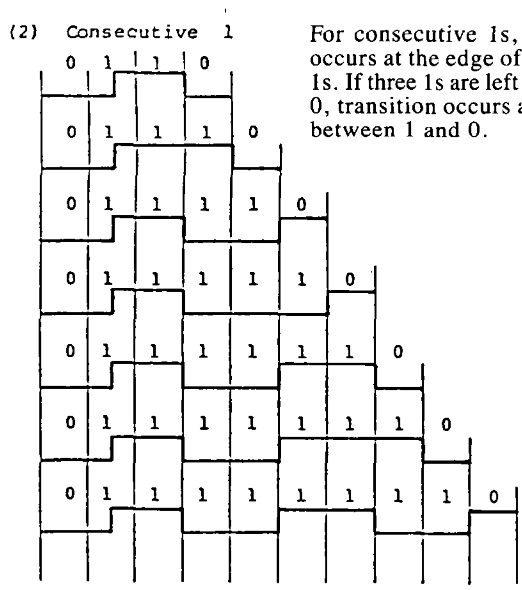


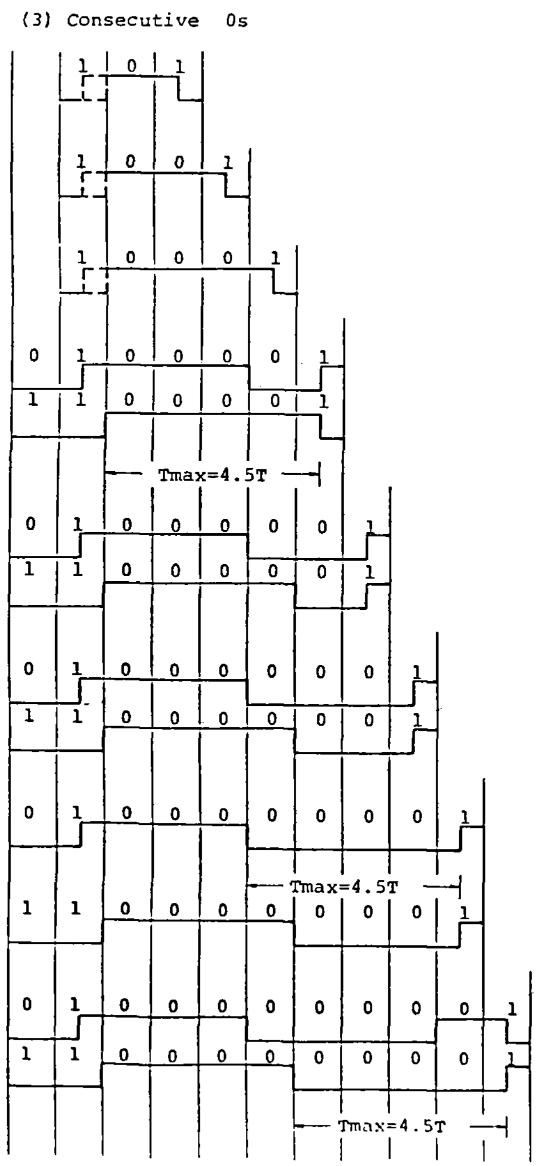
Fig. 9. Waveforms of newly developed channel codes.



Transition at the center of 1 bit cell.



For consecutive 1s, transition occurs at the edge of every two 1s. If three 1s are left just before 0, transition occurs at the edge between 1 and 0.



If the preceding transition exists at the boundary of a bit cell, no transition occurs for the succeeding 0s unless the transition at the boundary of the fourth and fifth 0s for at least five consecutive 0s.

If the preceding transition exists at the center of a bit cell, no transition occurs for the succeeding 0s unless the transitions at the boundary of the third and fourth 0s for at least four consecutive 0s.

Fig. 10. Coding rules of HDM-1.

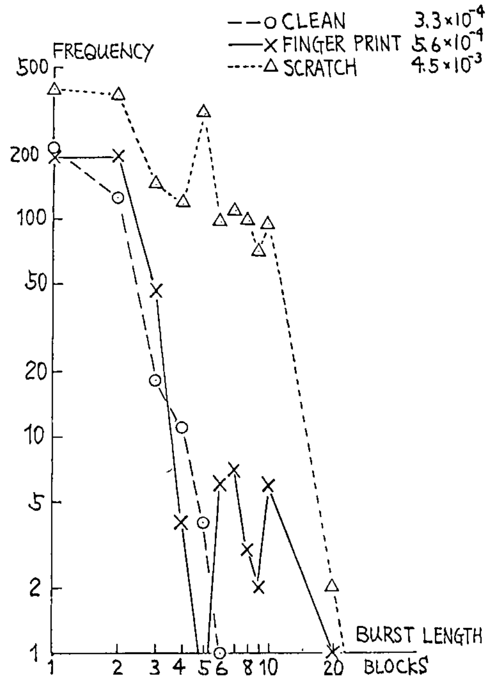
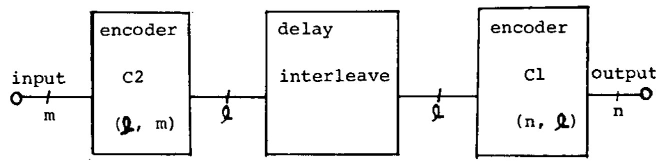


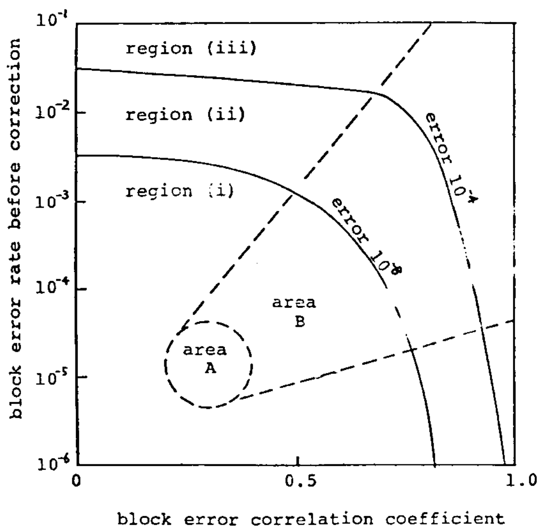
Fig. 11. Error distribution in optical disk system. Optical disk; 1.2 m/s; 2.352 Mbit/s; HDM-2; 1 block = 160 bits; f_s = 44.1 kHz.



DASH
C2 (7, 6) single-erasure correction
C1 (8, 7) single-erasure correction

Compact Disc (CIRC = cross interleave Reed–Solomon code)
C2 (28, 24) Reed–Solomon
C1 (32, 28) Reed–Solomon

Fig. 13. Principle of cross interleave method.



Region (i)	Good, Block error rate after correction $< 10^{-8}$
Region (ii)	Warning, $10^{-8} <$ prime $< 10^{-4}$
Region (iii)	Prohibited, $10^{-4} <$ prime
Area A	Best tuned condition
Area B	Deteriorated in studio environment

Fig. 12. Evaluation of DASH format.

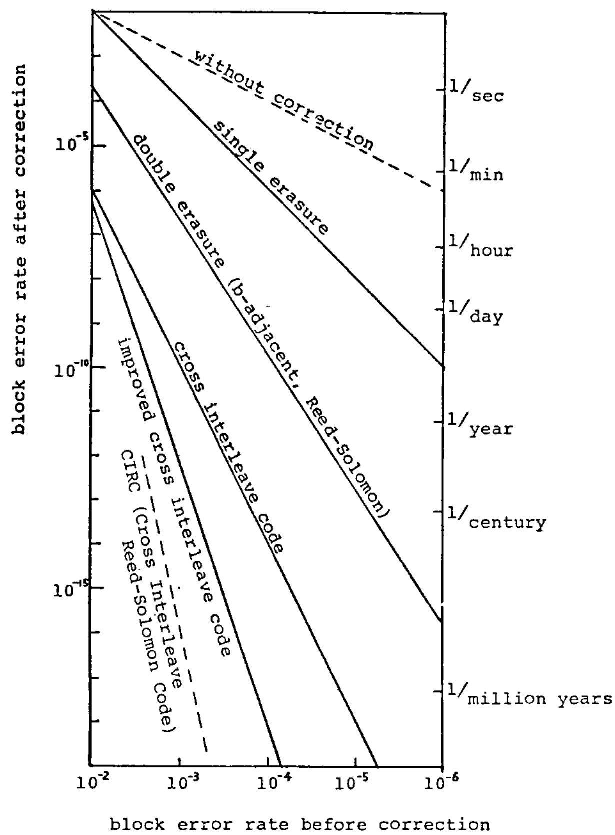


Fig. 14. Random-error correctability of various codes.

Table 4. Burst-error protection.

Format	Item	Correction (bits)	Concealment (bits) Good Concealment	Concealment (bits) Marginal Concealment
EIAJ	Rotary head, tape	4096	—	8 192
DASH	Stationary head, tape	8640	33 982	83 232
CD	Optical disk	3874	13 282	15 495

compressed form and read out at variable speed so that the editing point can be chosen in a manner similar to the shuttling of analog tape reels.

2) A small offset of two tape transports is adjusted in digital memory when synchonization is not perfect.

3) Data around the editing point are memorized to perform cross fading.

Editing in a multichannel recorder is somewhat different from that of two-channel systems. Even a simple punching in and out is relatively complex. As shown in Fig. 16, it is essentially a rerecording of the reproduced data and their cross fading with the input signal.

Special features are possible, as illustrated in Fig. 17, which is an assembly of multiple recordings on one channel in the digital domain, without using a digital mixing console.

Electronic editing using two or more multichannel recorders has more flexibility, as illustrated in Fig. 18. Here the editing points as well as the cross fading times can be chosen independently channel by channel.

Tape-splice editing is also possible in digital audio recorders, but complex signal-processing is necessary. Fig. 19 shows one of the methods with a large interleave between even and odd data, forming an effective overlapped area for cross fading at the splice.

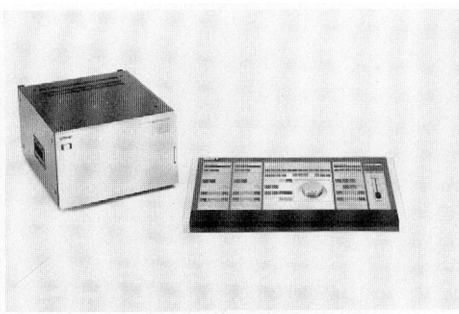


Fig. 15. A two-channel electronic editor.

2 COMPACT DISC SYSTEMS

An optical digital audio disk system was developed jointly by Sony and Philips [7], which has been marketed since the fall of 1982. The main specifications are given in Table 5. The same disk will be used by broadcasters who will appreciate its capability for quick random access.

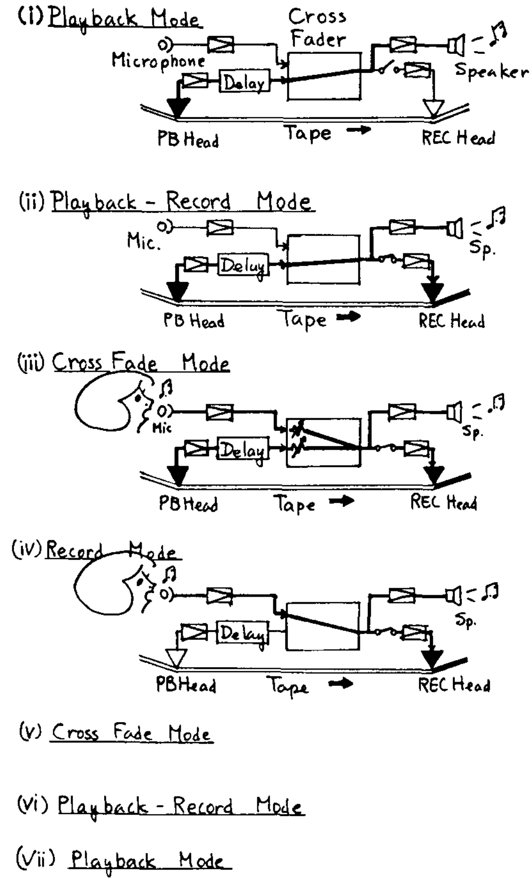


Fig. 16. Procedure for punching in and out.

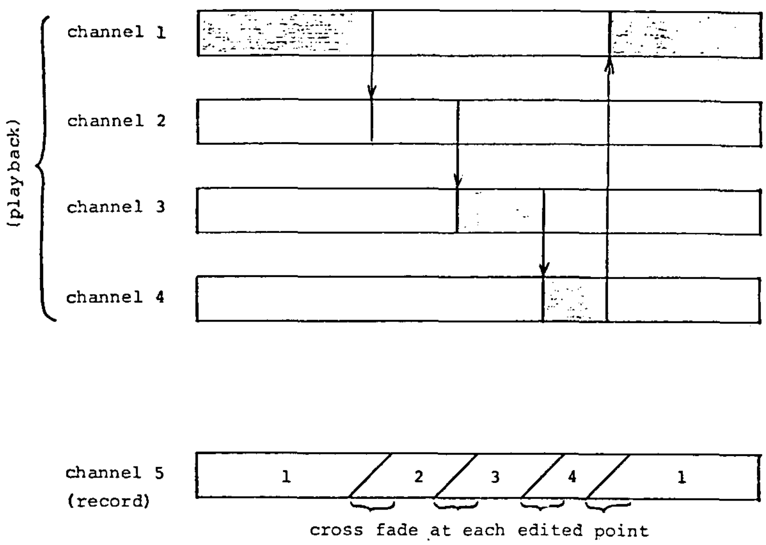


Fig. 17. Sequential punching.

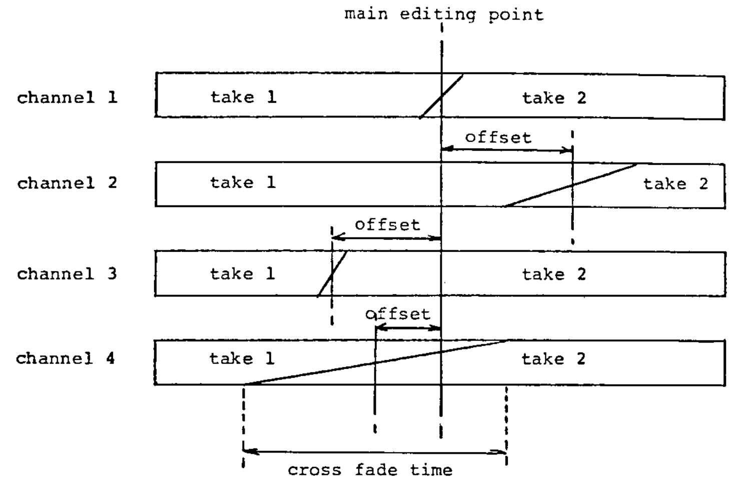


Fig. 18. Electronic editing with different cross fade times and offsets in each channel.

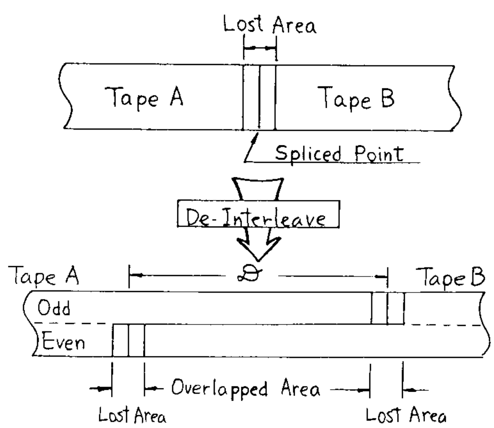


Fig. 19. Tape-splice editing.

Table 5. Specification of Compact Disc.

Playing time	75 min
Rotating speed	1.2–1.4 m/s (constant linear velocity)
Track pitch	1.6 μm
Disk diameter	120 mm
Disk thickness	1.2 mm
Center hole	15 mm
Signal surface	50–116ϕ mm (signal starts from inside)
Channel number	2
Quantization	16-bit linear per channel
Sampling rate	44.1 kHz
Data rate	2.0338 Mbit/s
Channel bit rate	4.3218 Mbit/s
Error protection	CIRC (cross interleave Reed–Solomon code) redundancy 25% (4/3)
Modulation	EFM (eight-to-fourteen modulation)

3 MULTICHANNEL RECORDER FOR PROFESSIONAL USE

The outline of an advanced format, the DASH is shown in Table 6 [8]. A 24-channel machine based on this format was developed [9]. Its size and appearance are shown in Fig. 20; rough specifications are given in Table 7. All the technology described is applied in this machine, and yet it is smaller than some analog recorders.

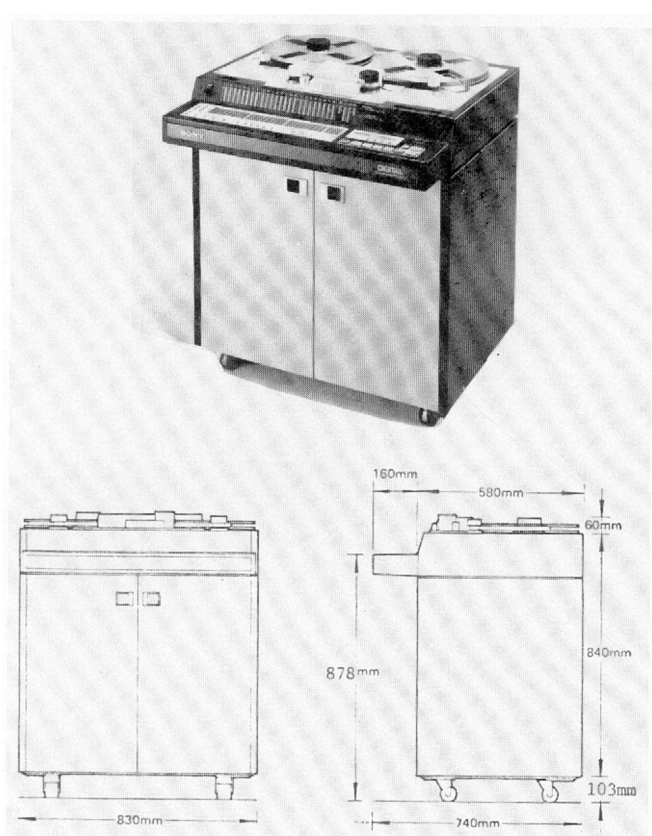


Fig. 20. Outside view of PCM-3324.

4 SATELLITE BROADCASTING

Signal-processing technology developed for digital audio recording is also applicable in satellite broadcasting. Fig. 21 shows a system that transmits 24 digital audio channels in a bandwidth equivalent to one video program [10]. The main specifications are given in Table 8.

5 LSI CIRCUITS DEVELOPED FOR CONSUMER APPLICATION

Fig. 22 shows three LSI circuits and a D/A converter developed for the CD system, and a player using them. Fig. 23 shows three other LSI circuits for signal processing in the EIAJ format [11] for a consumer digital audio recorder.

These LSI circuits reduce both costs and power con-

Table 6. Outline of the DASH format.

Version	Fast			Medium			Slow
Tape width (in)	¼	½	1	¼	½	1	¼
Digital channels	8	24	48	4	12	24	2
Tracks							
Digital	8	24	48	8	24	48	8
Analog	2	2	2	2	2	2	2
Time code	1	1	1	1	1	1	1
Control	1	1	1	1	1	1	1
Total	12	28	52	12	28	25	12
Sampling rate f_s (kHz)	48.0, 44.1, 32.0						
Tape speed (cm/s) at							
f_s = 48.0 kHz	76.20			38.10			19.05
f_s = 44.1 kHz	70.01			35.00			17.50
f_s = 32.0 kHz	50.80			25.40			12.70
Quantization	16-bit linear per channel						
Channel coding	HDM-1						
Error correction	Cross interleave, CRCC						
Redundancy	33% (error correction, detection, and synchronization)						
Bit length (μm)	0.6612						
Minimum wave length to be recorded (μm)	1.9837						
Packing density (bit/in)	38 400						

Table 7. Specification of PCM-3324.

Number of Channel (one track per channel)
- Digital audio 24
- Analog audio 2
- Time code 1
- Control 1
- Total 28

Tape speed, sampling rate
- 70.01 cm/s, 44.1 kHz, 76.20 cm/s, 48.0 kHz, with ±12.5% vernier
- (selectable at recording, automatic switching at playback)

Tape
- 0.5-in (12.7-mm) digital audio tape

Quantizaiton
- 16-bit linear per channel

Dynamic range
- More than 90 dB

Frequency response
- 20 Hz to 20 kHz $\left\{ \begin{matrix} +0.5 \\ -1.0 \end{matrix} \right\}$ dB

Total harmonic distortion
- Less than 0.05%

Wow and flutter
- Undetectable

Emphasis
- 50 μs/15μs (EIAJ format and Compact Disc compatible)
  - (on/off switchable (each channel) at recording and automatic switching at playback)

Format
- DASH-F (fast version of DASH)

Channel coding
- HDM-1

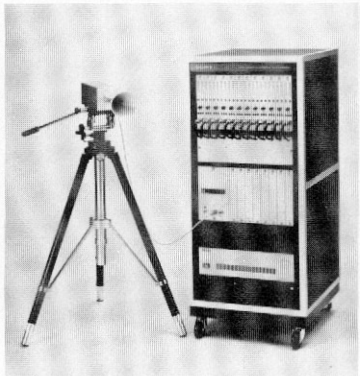
Error control
- Cross interleave code

Editing
1) Punch in/out
   - Cross fade time (variable), 1.45–372 ms (44.1 kHz)
2) Electronic editing
   - Editing point and cross fading time are set independently to each channel
   - Sequential punching in digital domain is possible
3) Tape splice editing
   - Cross fade time (fixed) 5.66 ms (44.1 kHz)
   - Tape cut at any point

Dimensions
- 830 (W) × 1003 (H) × 740 (D) mm

Weight
- 220 kg

Power consumption
- 2 kVA, 1.8 kW (peak)



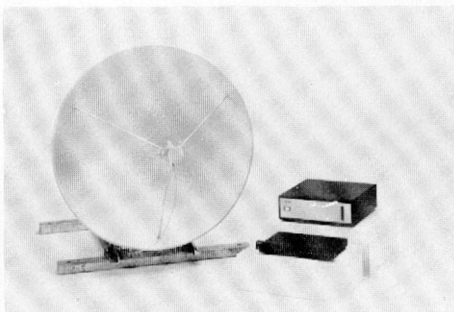


Fig. 21. Experimental system for satellite broadcast (Sony).

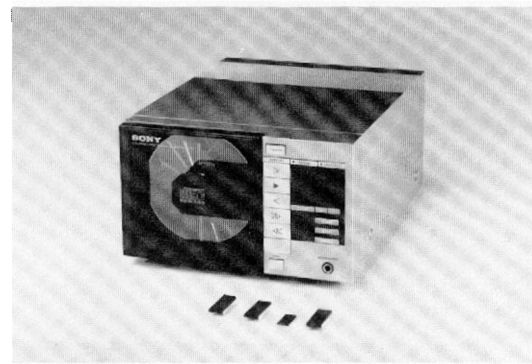


Fig. 22. LSI circuits for Compact Disc and player.

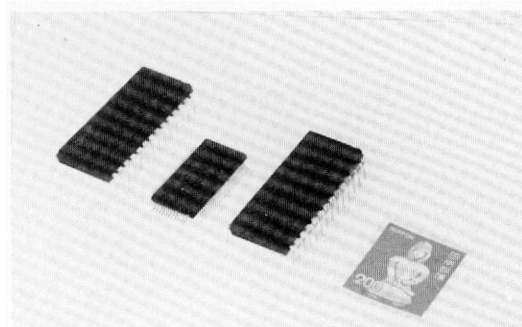


Fig. 23. LSI circuits for EIAJ format.

Table 8. Main specifications of experimental satellite broadcasting system.

Number of program channels	24 mono (or 12 stereo)
Sampling rate	32 kHz
Quantization	14-bit linear
Program frequency bandwidth	dc, ~15 kHz
Dynamic range	More than 80 dB
Error control	BCH + parity
Bit rate	18.432 Mbit/s
Modulation	Four-phase PSK

sumption and improve system reliability. It should be noted, however, that the huge investment made in their development is compensated only by mass production.

6 CONCLUSION

Some of the basic technologies recently developed in digital audio are described here. They support the second generation of digital audio systems, such as the Compact Disc, professional recorders, and satellite broadcasting.

As the applications expand, a more systematic approach is required to develop a total view of digital audio, from the production of music through its distribution (by way of tape, disk, and radio) to home entertainment equipment.

7 REFERENCES

[1] A. Kayanuma et al., "An Integrated 16-Bit A/D Converter for PCM Audio Systems," presented at the 1981 International Solid-State Conference, 1981 Feb.

[2] I. Hata et al., "An Integrated 16-Bit D/A Converter for PCM Audio Systems," Solid State Device Conf., Japanese Society of Applied Physics (1981 Aug.).

[3] T. Sekiya, H. Uchida, J. Sugita, H. Nakajima, and K. Odaka, "Digital Audio Compact Cassette Deck with Thin-Film Head," presented at the 71st Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 30, p. 362 (1982 May), preprint 1859.

[4] T. T. Doi, "Channel Codings for Digital Audio Recordings," presented at the 70th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 29, p. 936 (1981 Dec.), preprint 1856.

[5] T. T. Doi, K. Odaka, G. Fukuda, and S. Furukawa, "Cross-Interleave Code for Error Correction of Digital Audio Systems," presented at the 64th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 27, p. 1028 (1979 Dec.), preprint 1559.

[6] T. Ohtsuki, S. Kazami, M. Waiari, M. Tanaka, and T. T. Doi, "A Digital Audio Editor," presented at the 68th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 29, p. 358 (1981 May), preprint 1743.

[7] L. B. Vries, K. A. Immink, J. G. Nijboer, H. Hoeve, J. Timmermans, L. M. Driessen, T. T. Doi, K. Odaka, S. Furukawa, I. K. Iwamoto, Y. Sako, H. Agawa, and T. Itoh, "The Digital Compact Disc—Modulation and Error-Correction Schemes," presented at the 67th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 28, p. 931 (1980 Dec.), preprint 1674.

[8] T. T. Doi, Y. Tsuchiya, M. Tanaka, and N. Watanabe, "A Format for a Stationary-Head Digital Audio Recorder Which Covers a Wide Range of Applications," presented at the 67th Convention of the Audio Engi-

neering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 28, p. 931 (1980 Dec.), preprint 1677.

[9] T. T. Doi, "Design of Professional Digital Audio Recorder," presented at the 71st Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 30, p. 360 (1982 May), preprint 1885.

[10] H. Nakajima, A. Iga, H. Hideshima, T. Saito, T. Sato, and E. Fujita, "Satellite Broadcasting System for Digital Audio," presented at the 70th Convention of the Audio Engineering Society, *J. Audio Eng. Soc. (Abstracts)*, vol. 29, p. 941 (1981 Dec.), preprint 1855.

[11] EIAJ Tech. File STC-007, STC-008.

THE AUTHOR

Toshi Tada Doi received a B.E. degree in 1964 from the Tokyo Institute of Technology, Tokyo, Japan, and a Dr. Eng. degree in 1972 from Tohoku University, Sendai, Japan. Since 1964, he has been working for Sony Corporation, where he has been involved in various projects in the field of circuit theory, computer-aided design of integrated circuits and antennas. Since 1975, he has been head of the development team working on digital audio and has completed various digital audio systems both for consumer use and professional use. He contributed to establishing the EIAJ standard for domestic digital audio recorders, and developed the format for digital audio stationary-head (DASH) recorders for professional use. Dr. Doi also worked on the development of digital audio disk systems and contributed to establishing the format of the Compact Disc with the cooperation of Philips.

He was elected a fellow of the AES in 1980 May and a vice president of the International Region of the AES for 1982–1983. He received the Eduard Rhein award (Germany) in 1981 September and Michel de Coanda award (French) in 1982 March, both for the development of the Compact Disc systems. Dr. Doi is a member of the IECE, IEEE, AES, and ASJ.