The Origin of Digital Information Devices: The Silicon Audio and Its Family

AKIHIKO SUGIYAMA [1], AND MASAHIRO IWADARE[2]

This paper presents the origin of digital information devices, the Silicon Audio, and its family. The Silicon Audio is the digital counterpart of the Walkman and the ancestor of the iPod. It employs the MPEG / Audio Layer II algorithm for data compression, which was standardized by ISO (International Standardization Organization)/ IEC (International Electrotechnical Commission). A semiconductor memory card is equipped with to store the compressed signal. Since it has no mechanical movement, it is robust against shocks and vibrations that had been a serious problem for portable audio players. The background of the development, implementations, challenges toward a commercial product, and impact on audio players as well as personal information devices are discussed with its family including a video derivative, the Silicon View.

Keywords: Audio coding, Algorithm, MPEG, Semiconductor, Memory, Solid-state, Digital information device.

I. INTRODUCTION

The Walkman[1], which has become a byword for all portable audio players and is included in Oxford English Dictionary since 1986, was put in the market on July 1, 1979. With this product, people started to carry music outdoors. The history of portable information devices dates back to the Walkman which shipped 150 million units by 1995. The Walkman is a fully analog product. It has a cassette magnetic tape in it as the recording medium. With the dissemination of the Walkman in the 1980s, there were two big advances in electronics; semiconductor memories and data compression. The capacity of a single DRAM (dynamic access memory) chip increased from 64 Kbit in 1980 to 1Mbit in 1984, and to 64 Mbit in 1986. ITU-T (international telecommunications union, telecommunications sector), formerly, CCITT (comite consultatif international et telephonique), established a series of international standards for speech coding such as G.721 (1984) [1], G.722 (1986) [2], G.723 (1988) [3], G.726 (1990) [4], and G.728 (1992) [5] for efficient telecommunication. G.722 was developed for wideband speech signals to cover the 7 kHz bandwidth known as the wideband speech whereas G.721, G.723, G.726, and G.728 are for the telephone bandwidth up to 4 kHz. It was a sign that the wideband speech with a 7 kHz bandwidth and the CD (compact disc) quality audio signal with a 20 kHz bandwidth would be in the center of the market.

Advances in hifi audio coding algorithms in early 1990s made it possible to compress 20 kHz audio data into one

Table 1. Compression ratio of audio coding algorithms available in early 1990s.

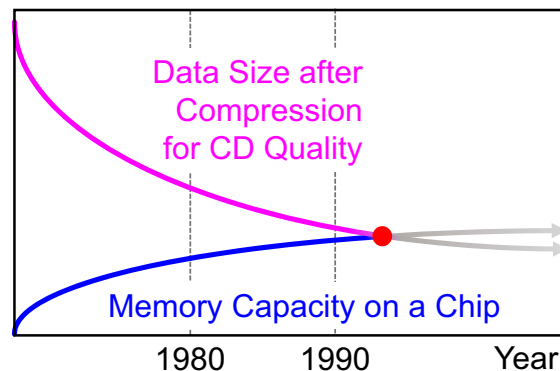| Algorithm	Compression Ratio	Coding
AC-2	1/6	Transform Coding
ATRAC	1/5	Hybrid Coding
PASC	1/4	Subband Coding
MPEG	$1/4 \sim 1/12$	Hybrid Coding

Fig. 1. Capacity on a chip meets data size after compression.

fourth to one twelfth of the original without noticeable distortion. These algorithms include AC-2 for inter-studio transmissions [6], ATRAC (adaptive transform acoustic coding) for MD (mini disc) [7], PASC (precision adaptive subband coding) for DCC (digital compact cassette) [8], and MPEG-1/Audio algorithms [9] for CD-ROM (read-only memory). The compression ratio of these algorithms are compared in Tab. 1. They had been developed to decrease the size of the audio signal in storage or increase efficiency in the transmission channel. It was important because information to be stored or transmitted for a longer listening time and/or a better presentation with a wider bandwidth was sharply increasing in those days.

[1]Data Science Research Laboratories, NEC Corporation, 1753 Shimonumabe, Nakahara-ku, Kawasaki 211–8666, Japan
[2]Intellectual Property Management Division, NEC Corporation, 7-1 Shiba 5-chome, Minato-ku, Tokyo 108–0014, Japan

Corresponding author: Akihiko Sugiyama
Email: a.sugiyama@ieee.org
[1]*Walkman* is a registered trademark of Sony Corporation.

Traditionally, the compressed data had been stored in an optical or a magnetic device such as optical/magnetic discs and magnetic tapes. However, requirements for audio devices needed to include longer recording time and quick random access. In addition, compact size and light weight for easy handling as well as shockproof are also important for portable devices. In view of these requirements, semiconductor memories could be a new candidate. Specifically, it was of general interest to store the compressed data onto a semiconductor-memory card. A portable audio device with a semiconductor memory card was a natural consequence of an intersection between the ever-increasing memory capacity on a single chip and the ever-decreasing number of bits to express a practical length of audio signal with the state-of-the-art audio coding algorithm as shown in Fig. 1. It was December 1st, 1994, when the Silicon Audio made a debut as the world's first all solid-state portable audio player [10]. It is the ancestor of the iPod as well as the origin of a variety of digital information devices to appear in the following years.

This paper presents the origin of digital information devices, the Silicon Audio, and its family. The Silicon Audio naturally leads to its video derivative, the Silicon View [11], which paved the way to a wider variety of digital information devices including "the iPod with video" [12]. In the following section, conventional audio players are reviewed from a viewpoint of their shortcomings. The new audio-data encoding/decoding system with semiconductor storage is introduced in Section III with its challenges to a successful commercial product in Section IV. Section V presents the Silicon Audio family as promising future digital information devices in those days. Finally, in Section VI, impact of the Silicon Audio on audio players and personal information devices is discussed.

II. CONVENTIONAL AUDIO PLAYERS

Portable cassette players had been widely used for personal entertainment since Sony released the Walkman in 1979. They use a signal processed and recorded by an analog system, thus, suffer from noise. In late 1980s, digital audio tape (DAT) recorders were introduced in the market which was realized by digital circuit. However, a common disadvantage of these systems is lack of random-access capability because they are both equipped with a magnetic tape as storage. The tape has to be rewound for playback and fast-forwarded to skip some part of the contents. Compact-disc (CD) Players were realized based on all-digital technologies in early 1980s. It supports 20 kHz for the signal as the bandwidth and provides random access. Nevertheless, the mechanical movement in the player causes undesirable skips in the reproduced sound by shakes and vibrations, which degrade sound quality. In addition, the weight and the size were yet other drawbacks of such systems, not to mention power consumption and reliability. An audio player with semiconductor memory had the potential to solve all these problems.

Fig. 2. The Silicon Audio and its promising scenarios.

III. THE SILICON AUDIO

A) System Features

The Silicon Audio[2] [13–15] consists of three parts: an encoder to compress the incoming source signal, a semiconductor card memory to store the compressed data, and a decoder for decompression of the stored audio data to reproduce the original audio signal. The MPEG Layer II algorithm is employed for encoding/decoding. Thanks to this algorithm, a 20-kHz bandwidth, which is equivalent to the CD quality, is supported with minimum audible distortion. The Silicon Audio has no annoying skip in the reproduced sound even with shakes and vibrations because it has no mechanically moving parts. As additional advantages, it is light-weight, compact, and comes with low power consumption thanks to all solid-state implementation, thus, useful in various scenarios with jumps, shocks, and vibrations as illustrated in Fig. 2.

B) Coding Algorithm

The coding algorithm is a key to the overall quality of the system. A good algorithm achieves a higher compression rate and requires less memory capacity for the same source signal. MPEG-1/Audio algorithms achieved better compression rate than other algorithms as demonstrated by subjective evaluations [16, 17]. In addition, they have possibilities to require less cost than other *de facto* standards for implementation because MPEG is an international standard and was expected to be employed by a larger number of products. The MPEG-1/Audio algorithms have a layered structure with Layers I, II, and III. Among these algorithms, the Layer II algorithm is a good compromise between the compression ratio and the quality of the reproduced signal. Therefore, the Silicon Audio employed the MPEG-1/Audio Layer II algorithm.

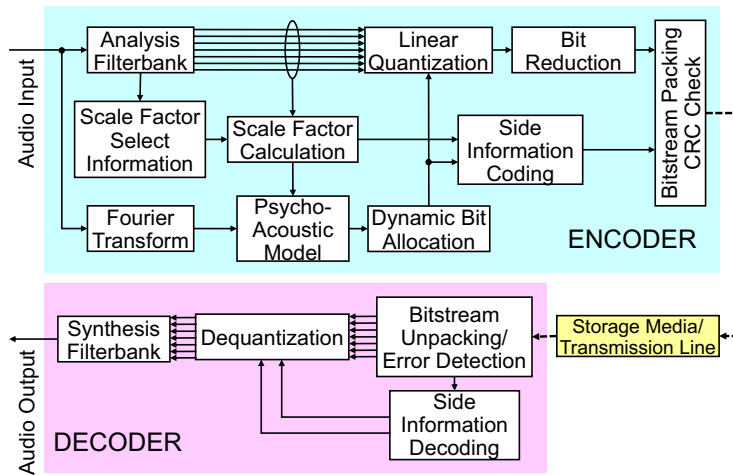[2]*Silicon Audio* is a registered trademark of NEC Corporation.

Fig. 3. Encoding and decoding process of MPEG I, Layer II.

Fig. 4. Memory capacity vs. recording time (96, 128, 192 kbit/s/ch). Vertical dashed lines represent some typical capacities in mega bytes wheras the abscissa is expressed in mega bits.

Fig. 5. Subjective evaluation results of MPEG-1 Audio.

The overall encoding and decoding processes of the Layer II algorithm are illustrated in Fig. 3. The incoming linear-PCM (pulse code modulation) data are split into 32 subbands by a polyphase analysis filter bank. Scale factor select information is defined from subband samples and scale factors are calculated. The fast-Fourier-transformed data of the same input signal as to the analysis filter bank are used to calculate masking thresholds based on a psychoacoustic model. The obtained thresholds are used for dynamic bit allocation. Subband samples, quantized with the obtained bit allocation, are formatted with the header and side information into the bitstream.

The decoding process could be described by tracing back the encoding process. Before bitstream unpacking, errors are detected by examining the CRC (cyclic redundancy code) check bits. The error correction method is not defined by the international standard and left open for different user requirements. The compressed data are unpacked from the bitstream into the header, side information and quantized samples. The quantized subband samples are then inversely quantized by the de-quantization process with the assigned number of bits. The synthesis filter bank combines the de-quantized 32-subband samples into a single full-band signal to reproduce the output signal.

C) Bit Rate and Recording Time

The recording time T in seconds is defined as a function of the memory capacity M_c, the bitrate per channel B_r, and the number of channels N_c as

$$T = \frac{M_c}{B_r \times N_c} = \frac{8m_c}{B_r \times N_c}. \tag{1}$$

M_c, m_c and B_r are expressed in bits, bytes and bit/s , respectively. N_c is equal to 2 for stereo presentation and 1 for mono. Relations between the memory capacity and the recording time are summarized in Fig. 4 for different bitrates.
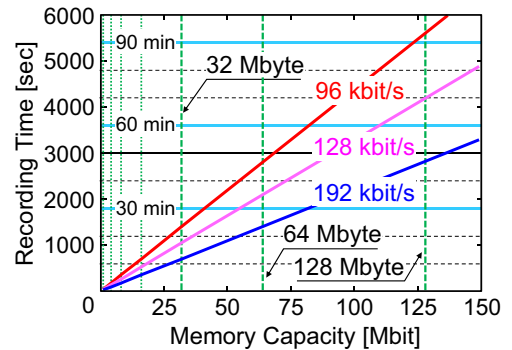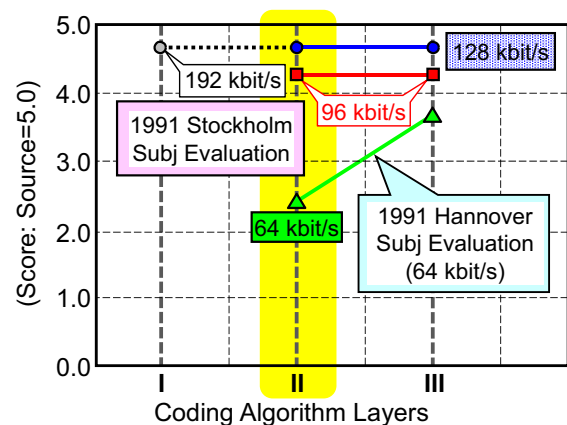
The abscissa and ordinate are the memory capacity in megabits and the recording time in seconds, respectively. Each horizontal dashed line corresponds to 10-minute recording and each vertical dashed line is used to show a memory capacity expressed by a power of two such as 32, 64, 128, etc. in Mbyte. The three slant solid lines are the memory capacity vs. recording time at 96, 128, and 192 kbit/s/ch. Commercially available flash memory cards had as much as 40 Mbyte (equivalent to 320 Mbit) as the capacity in mid 1990s. It means that almost 30-minute recording was possible at a rate of 96 kbit/s/ch. For 60-minute recording, which could be considered as one of the minimum recording times imposed by consumer requirements, 640 megabits (80 Mbytes) are necessary as the memory capacity. It was just twice as large a capacity as the current one and did not take long to be achieved by evolving card-memory technologies in those days.

The target bitrates for the Layer II algorithm are 96 and 128 kbit/s/ch. Figure 5 illustrates the results of the two official subjective tests in 1991 [16, 17]. The Layer II provides

Fig. 6. Three generations of the Silicon Audio developed in early 1994 [13], late 1994 [14], and 1997 [15] (Left to right).

AMP: Amplifier   BS: Bitstream
DAC: Digital-to-Analog Converter
DAI: Digital Audio Interface
PSCNV: Parallel-to-Serial Converter
DSPSEL: Buffer Control Signal for DSP
RESET: Reset Signal INT: Interrupt Signal
SO: Serial Output BUF: Buffer INTF: Interface
SOCK: Serial-Output Clock
LRCLK: Clock Signal for L/R Switching

Fig. 8. Blockdiagram of the Silicon Audio player (first generation).
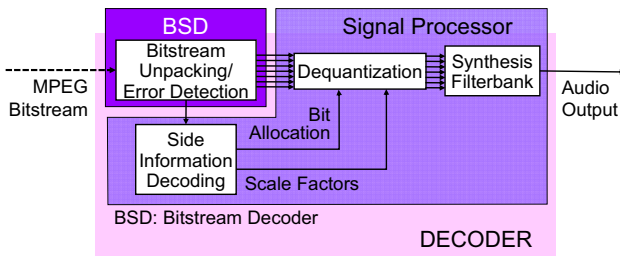

Fig. 7. Task distribution between BSD and SP.

transparent quality at a bitrate of 128 kbit/s/ch, and some reasonably good quality even at 96 kbit/s/ch. As the Silicon Audio is mainly targeted at general consumers, the incoming data are encoded generally at a bitrate of 96 kbit/s/ch for a reasonable compromise between the recording time and sound quality. However, the bitrate is user-selective from among 192, 128, 96, and 64 kbit/s/ch.

D) Implementation

The first generation in early 1994 was a result of rapid prototyping. The analog circuit for AGC (automatic gain control) and ADC (analog-to-digital conversion) as well as the digital circuit for decoding were separately power-supplied at 15V and 5V, respectively, for easy implementation. There were no internal batteries so that the power had to be supplied from outside. The second generation prototype in late 1994 was implemented with an MPEG-1/Audio decoder LSI and powered by four 1.2V nickel metal hydride (NiMH) batteries used for commercial portable audio players.

It was not possible to accommodate both encoding and decoding process in a compact portable player. There was no commercial encoder or decoder chip available in the market in early 1990s. Digital signal processors (DSPs) were not powerful enough to perform both encoding and decoding in realtime. Therefore, play-back-only part with a card memory was implemented as a prototype system. Encoding was carried out either by a breadboard system
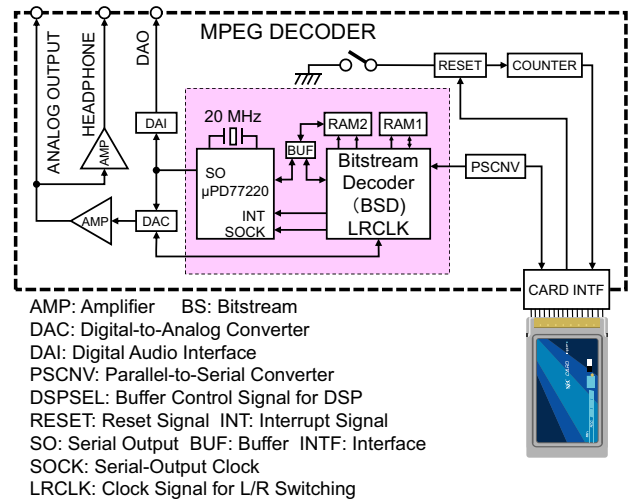
or a software encoder simulator. Figure 6 depicts a picture of the Silicon Audio developed in early 1994 [13], late 1994 [14], and 1997 [15].

1) First Generation: DSP+Gate Array Implementation

The decoder was implemented by two chips; an NEC's 24-bit fixed point signal processor (SP) μPD77220 [18], for most of signal processing, and a gatearray chip, called Bitstream Decoder (BSD), for timing generation, bitstream unpacking and error detection by CRC check. Task distribution between BSD and the SP is depicted in Fig. 7 with an overlapped functional diagram of the Layer II decoder.

Shown in Fig. 8 is a schematic diagram of the Silicon Audio decoder of the first generation. Encoded data, which are stored on a card memory, are retrieved and converted to serial data. The serial-data bitstream, which is fed to BSD, is first used for synchronization and then, written in RAM1. The data stored in RAM1 are read out, unpacked in BSD, and stored in RAM2. Whenever the synchronization pattern is detected in BSD, a frame synchronization signal is provided with the SP as the reset signal. The interrupt signal and the serial clock signal, both used for output timing, are also provided by BSD.

The SP has an access to RAM2 and obtains data therefrom. It is devoted to most of the decoding process such as side-information decoding, de-quantization of the subband samples and synthesis of the full-band signal by a filter bank. The SP is operated in the master mode at a machine cycle of 100 ns. The decoded audio sequence is supplied to the digital-to-analog converter (DAC), which both the left and the right channel signals share. Switching between both channels is controlled by a signal generated by BSD. The output signals are then supplied to the output terminals via an analog stereo amplifier. A digital audio interface (DAI) and a terminal for headphones with an additional analog amplifier are also accommodated.

All these LSIs and discrete parts as well as a memory card are accommodated in a package with a dimension of
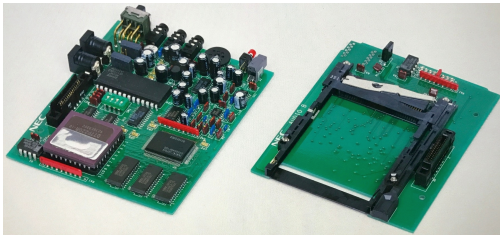
Fig. 9. Printed-circuit board of the Silicon Audio player (first generation). Top view (left) and bottom view (right). The SP and BSD are located at the close end of the print-circuit board.

Fig. 10. Blockdiagram of the Silicon Audio player (second generation).

Table 2. Specifications of Silicon Audio Player (first generation).

Algorithm	MPEG-1/Audio Layer II
Bitrate	96 kbit/s/ch
Implementation	1 DSP+G/A, 24-bil Fixed Point
Dimension	$127 \times 98 \times 33$ mm
Recording Time	12 min / 16 Mbyte

Table 3. Features of the Decoder LSI

Algorithm	MPEG-1/Audio Layer I/II
	MPEG-2/Audio LSF Layer I/II
Bitrate:	All bitrates except free format
Sampling Frequency:	16, 22.05, 24, 32, 44.1, 48 kHz
Mode:	Stereo, Dual-Channel,
	Joint Stereo, Mono
Input Bitstream:	Serial/Parallel Selective
Max. Operating Freq.:	24.576 MHz
Package:	100-pin QFP
Power Supply:	5 V
Power Consumption:	350 mW

Fig. 11. Print-circuit board of the Silicon Audio player (second generation). Top view (left) and bottom view (right). The decoder LSI is located at the close end of the left picture.

$127 \times 98 \times 33$ mm. Figure 9 depicts the print-circuit board of the Silicon Audio player implemented with a DSP and a gate array chip. The specifications are summarized in Tab. 2.

2) Second Generation: Single Chip Implementation

The second generation was implemented with only a single chip MPEG-1/Audio decoder LSI. Features of the decoder LSI are summarized in Tab. 3. The MPEG-1/Audio decoder LSI [19, 20] conforms to Layers I and II algorithms in ISO/IEC 11172-3, known as MPEG-1 Audio. To achieve low power-consumption and single-chip realization, a hardwired approach was employed. The LSI also provides functions to decode Layers I and II LSF (low sampling-frequency) bitstreams of ISO/IEC 13818-3 [21]. The filter bank is designed as a cosine-modulated filter bank [22]. By using the symmetry of the cosine function and a fast DCT (discrete cosine transform) [23], it finally requires 30% multiplications and 37% additions of the original calculation [24].

Shown in Fig. 10 is a schematic diagram of the Silicon Audio player, the second generation. Encoded data,
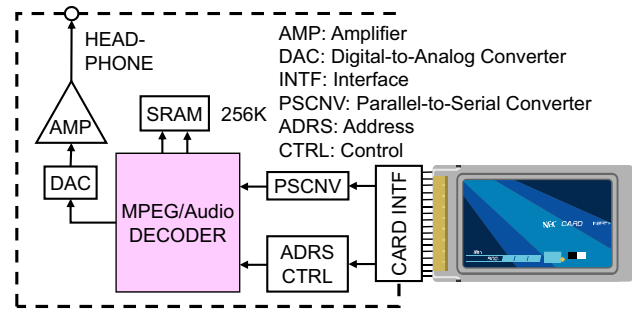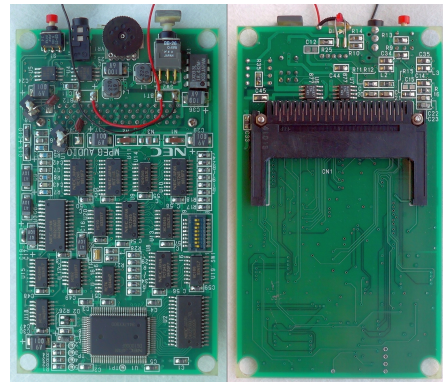
which are stored in a card memory, are retrieved and converted to serial data. The read address of the card memory is controlled by ADRS CTRL. The generated address is generally monotonically increasing, however, it is easy to jump to a new address instantly by a control command. The serial bitstream is fed to the decoder LSI. All the decoding processes such as synchronization, bitstream unpacking, side-information decoding, de-quantization of the subband samples and synthesis of the full-band signal by a filter bank are performed in the single-chip LSI. The decoded audio sequence is supplied to a DAC, which both the left and the right channel signals share. The output signals are then supplied to the output terminals via an analog stereo amplifier.

The decoder LSI and peripheral components as well as a memory card and four NiMH batteries are accommodated in a package with a dimension of $130 \times 78 \times 22$ mm. Figure 11 depicts the print-circuit board (PCB) of the Silicon Audio player. The audio decoder LSI is located at the close end of the left picture. The specifications of the Silicon Audio player (second generation) are summarized in Tab. 4.

Table 4. Specifications of the Silicon Audio Player (second generation)

Algorithm	MPEG-1/Audio Layer I/II
Bitrate	All bitrates except free format
	96 kbit/s/ch (Recommended)
Implementation	μPD60312 + Peripherals
Dimension	$131 \times 78 \times 22$ mm
Recording Time	24 min / 32 Mbyte (256 Mbit)
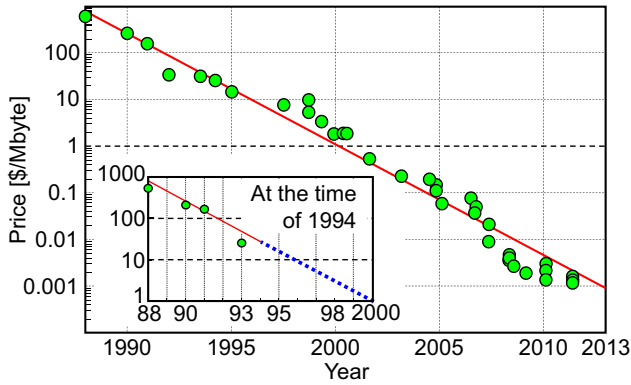
Fig. 12. Memory price in 1990s and 2000s [25–27].

IV. CHALLENGES TOWARD A SUCCESSFUL COMMERCIAL PRODUCT

The Silicon Audio was a prototype which had challenges to overcome to make a successful commercial product. They are the memory price, copyright protection, and contents distribution, to name a few. In mid 1990s, the memory price was not sufficiently low for daily use. In 1993, 1 Mbyte flash EEPROM (electrically erasable programmable read-only memory) and 1 Mbyte un-rewritable mask ROM cost $30 and $0.625, which were expected, by extrapolation, to go down as low as $1 and $0.021 per Mbyte by 2000 [25, 26]. To record 60 minutes of stereo sound, the necessary EEPROM capacity from (1) is 82 Mbytes which would still cost U$82 in 2000 if the prediction based on the inset [26] in Fig. 12 were correct.

Now, the history tells us that the memory price has been reduced continuously [27] as shown in Fig. 12. With the decline of the memory price, small companies in Korea and Singapore such as Saehan Information Systems, Diamond Multimedia, DigitalCast (Rio), ReignCom (iRiver), and Creative, released portable audio players in the market in late 1990s and Japanese electronics manufacturers such as Sony, Aiwa, Toshiba, Matsushita (Panasonic), and Sanyo followed in 2000s. The mpman MP F-10 by Saehan was released in 1998 as the first commercial all solid-state audio player. It had 32 or 64 MB internal memory to keep 30 or 60-minute music encoded at a bitrate of 128 bit/s, stereo. Because most of them employed MPEG-1/Audio Layer III, also known as MP3, as the coding algorithm, these players are called MP3 players. The AudioKey released in 2003 by Packard Bell, an NEC subsidiary, is one of those products.

The signature portable audio player by Apple, the iPod[3], was released in 2001. Even though it was one year after the EEPROM price/Mbyte had reached $1 as predicted in Fig. 12, the iPod was equipped with a hard disk drive (HDD) to provide sufficient memory capacity at a reasonable cost. All solid-state audio players by Apple had not been available until the iPod nano was released in 2005.

There was no single popular copyright protection method, thus, each product used its own method in the 1990s. It was because each small manufacturer could not negotiate with multiple content providers in parallel for a comprehensive contract to cover wide range of contents. Users need to collect or encode audio data by themselves piece by piece. Some of consumer electronics companies such as Sony and JVC were bigger than forerunners thus stronger in contract negotiation. However, they had their subsidiary content providers. Those with a subsidiary content provider were reluctant to access content providers outside the family to cover a wider range of contents which are more attractive for users. They spent time and effort to cooperate with each other to achieve a comprehensive contract with content providers. However, it was not easy for them to agree on a common digital rights management method. Apple did not try too hard to agree on a standard digital rights management method with other companies to avoid conflict with others. As a result, Apple was the first to succeed in a comprehensive contract with multiple content providers to implement one-stop shopping system for audio data. With its iTunes Music Store (currently, iTunes Store) established in 2003, Apple has a share of approximately 80% in 2012 in the music player market.

V. THE SILICON AUDIO FAMILY FOR OTHER APPLICATIONS

There are a wide range of potential applications which could be extended from the Silicon Audio. This is because card memories can store a variety of data such as sound, speech, text, still images and even movies. A direct extension from audio to video led to the Silicon View [28, 29].
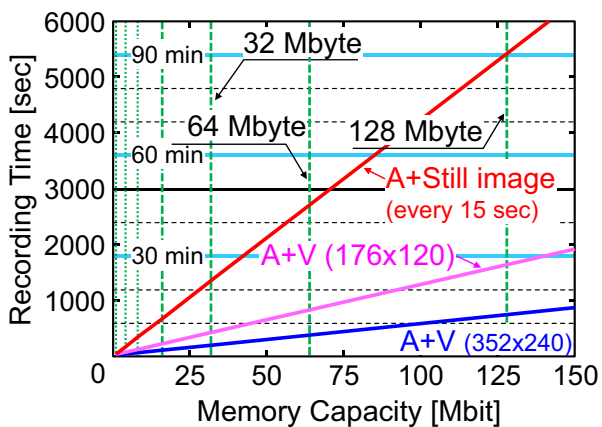
A) The Silicon View

The Silicon View[4] is a video extension of the Silicon Audio, which was revealed on October 12th, 1995. [11] It reads MPEG-1 Video [30] data from a semiconductor memory card, decodes them, and displays the decoded data in an liquid-crystal display (LCD) in addition to the Silicon Audio functions. The Silicon View was developed in two models; a portable model with a 5.5-inch LCD for better video presentation and a hand-held model with a 2.5-inch LCD for the minimum possible size and weight. Their specifications are summarized in Tab. 5.

A recommended bitrate for the compressed bitstream is about 1.4 Mbit/s or 10 Mbyte/min for the video with a

---

[3]*iPod, iPhone, iPad, iTunes Music Store,* and *iTunes Store* are registered trademarks of Apple Inc.

Table 5. Specifications of Silicon View players (portable and hand-held model).

| Version	Portable model	Hand-held model
Algorithm	MPEG-1 Video	Same as left
Number of pixels	352×240 max.	Same as left
Implementation	$\mu PD61010$ + Peripherals	Same as left
LCD Size	5.5 inch	2.5 inch
LCD Panel	TFT (thin film transistor)	TN (twisted nematic)
LCD Resolution	320×240 pixel	312×230 pixel
Pow. Consump.	5 V 800 mA 9.5 V 700 mA	5 V 600 mA 3.6 V 300 mA
Power Supply	External AC Adapt- ers	Lithium Ion Battery 3.6 V 1300 mAh
Dimension	$222 \times 125 \times 62$ mm	$146 \times 76 \times 37$ mm
Weight	1000 g	295 g with battery
Recording Time	4 min / 40 Mbyte	Same as left

Fig. 13. Memory Capacity vs. Recording Time for Silicon View with three modes: Audio + still images (every 15 sec), Audio + Video (176×120 pixels), and Audio + Video (352×240 pixels).

resolution of 352×240 pixels with a 44.1 kHz sampled stereo audio signal. The compression ratio is $1/23$. When the resolution is reduced to 176×120 pixels, the bitrate is one half with twice as long recording time as for the original resolution. The Silicon View player has a video zooming function which automatically displays the video program larger by horizontal and vertical pixel interpolation when the video input has a resolution smaller than 176×120 pixels. Use of images instead of moving pictures is possible, which enables approximately a tenfold playback time. Figure 13 illustrates relations between the memory capacity and the recording time for three different modes.

The Silicon View players consist of a single-chip MPEG-1 Audio/Video (AV) decoder LSI, $\mu PD61010$ [31], an address control unit (ADRS CTRL), a timing control

---

^{4}*Silicon View* is a registered trademarks of NEC Corporation.

Table 6. Features of the AV Decoder LSI

Video decoding	352×240@30 Hz, 352×288@25 Hz high-resolution still image (704×480 or 704×576 pixels)
Audio decoding	Stereo at 32, 44.1, or 48 kHz sampling
Max. Operating Freq.:	27 MHz
Package:	160-pin PQFP
Power Supply:	5 V
Power Consumption:	1.7 W
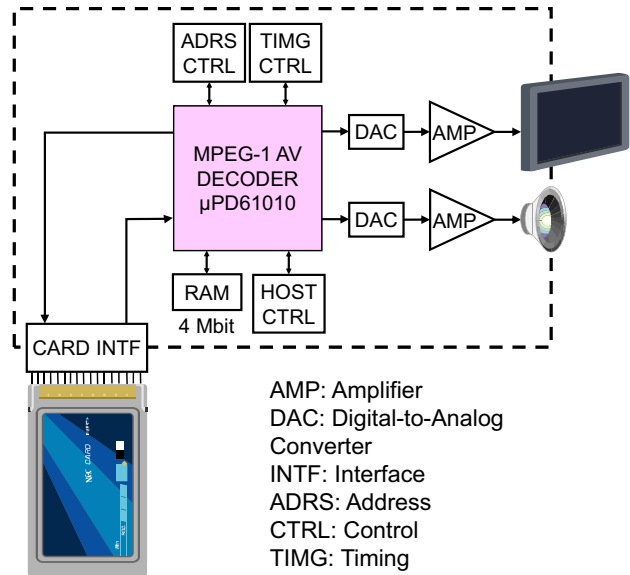
Fig. 14. Blockdiagram of the Silicon View players.

unit (TIMG CTRL), a 4-Mbit RAM, a host control unit (HOST CTRL), a pair of a DAC and an analog amplifiers for the output AV signals, an LCD, and a loudspeaker. Features of the AV decoder LSI and a blockdiagram of the Silicon View players are shown in Tab. 6 and Fig. 14.

The decoder LSI and the RAM receive an encoded bitstream from the card memory, decode it, and provide the DACs with a digital RGB (red-green-blue) video and a digital audio signal. A timing signal for decoding is generated by the TIMG CTRL unit. The analog AV signals from the DACs are amplified and drives the LCD and the loudspeaker. Addressing for the card memory is controlled by the ADRS CTRL unit which also takes care of the switching signal among "play," "stop," "pause," and "program jump." The HOST CTRL unit keeps an eye on the LSI status and resets the whole system upon an error. ADRS CTRL unit, HOST CTRL unit and TIMG CTRL unit are implemented by a single-chip microprocessor and a field programmable gate array (FPGA). The main decoding PCB for the hand-held model is 75×105 mm. The hand-held model of the Silicon View player operates for 40 minutes with a built-in rechargeable lithium ion battery. The portable model and the hand-held model of the Silicon View player are depicted in Fig. 15.

Fig. 15. Silicon View players: Portable model (left) and hand-held model (right).

Table 7. Specifications of AV Silicon Guide.

Algorithm	MPEG-1 Video
Number of pixels	320×240 dot
LCD panel	TFT (thin film transistor), 5.5 inch
Power Supply	Lithium ion battery, 7.2 V 2800 mAh
Operating time	90 min.
Dimension	$128 \times 200 \times 55$ mm
Weight	900 g (with battery)

Fig. 16. The Silicon Guide: AV model (left) and audio model (right).

B) The Silicon Guide

A more specific extension of the Silicon Audio and/or Silicon View is a handheld guide system which can be carried at an exhibit, a sightseeing spot, or in a museum and tells the user details of the displayed object. The Silicon Guide is such a system. The audio-only model is easier to carry and the audio explanation complements the displayed information. The AV model [32] is heavier, however, provides more detailed information in a more effective way with the mixed media. Multiple programs are played back, each of which contain AV data specific to the display. It has the following functions:

(i) Playback of the programs in the specified order.
(ii) Automatic hold after each program.
(iii) Repeat of the previous program upon user's direction.
(iv) Quick search for a desired program.

The Silicon Guide provides control of program-playback order and quick random access, which are inevitable for these functions. The specifications of the AV Silicon Guide are summarized in Tab. 7. Compared to the portable model of the Silicon View, the AV Silicon Guide operates on a single lithium ion battery for up to 90 minutes. Figure 16 depicts the AV model and the audio model of the Silicon Guide.

C) The Shopping Navigation

A possible origin of today's AV streaming terminals attached to a store shelf is the Shopping Navigation [33] which displays detailed information of a suggested product, the most popular product, or an event based on an

Fig. 17. The Shopping Navigation.

Table 8. Specifications of Shopping Navigation.

Algorithm	MPEG-1 Video
Number of pixels	720×220 dot
LCD panel	Twisted Nematic (TN), 4 inch
Power Supply	5 V 1 A, 12 V 0.35 A
Dimension	$145 \times 150 \times 45$ mm
Weight	420 g (without power supply)
Recording time	1 min / 8 Mbyte (video mode)

instruction by the user. The Shopping Navigation, depicted in Fig. 17, keeps playing back the main program with multiple selection indices such as A, B, and C until there is an instruction by the user to change the program. The user can press one of those labeled buttons near the display for instruction. Depending on which of A, B, or C button is pressed, a different program will be played back. Table 8 shows the specifications of the Shopping Navigation.

VI. IMPACT OF THE SILICON AUDIO ON AUDIO PLAYERS AND DIGITAL INFORMATION DEVICES

The Silicon Audio had an unparalleled impact and was reported in major newspapers and magazines in Japan, US, and UK such as the Asahi [34], the Nikkei [35], the Financial Times [36], the International Herald Tribune [37], Time [38], and the Future Music [39]. The Silicon Audio opened a new way to solid-state audio players. A decoder (and an encoder at a later time) is implemented by peripherals and an LSI chip which anybody can purchase from a semiconductor company. The necessary number of parts was significantly reduced. Such simplification is clearly demonstrated by comparing Figs. 8 and 10. Place a small number of electrical parts on the print circuit board and float it on the surface of a solder pool. That is the whole assembly process. No integral process for discrete parts based on long and extended experience in assembly is needed. Anybody can start manufacturing business. As a result, small companies as mentioned in Section IV newly started manufacturing audio players. It was a revolution in manufacturing which invited new companies to join and resulted in market expansion in the early stage. It is interesting to see the same thing happening in car industry today. Electric vehicles are like the Silicon Audio in comparison with conventional cars powered by fossil fuels such as gas and diesel fuel.

The Silicon Audio has made significant impact upon personal information devices leading to today and future as illustrated in Fig. 18. Today's most successful portable audio player, the iPod, was put in the market in 2001. This iPod player employed MPEG-1/Audio Layer III also known as MP3. It took three years until the iPod employed MPEG-4/Audio AAC (advanced audio coding) [40] that is today's most widely used MPEG/Audio standard. It should be noted that the iPod still used an HDD as the memory device. Next year in 2005, Apple released the iPod nano that employed semiconductor memory therein.

The next notable step was release of the iPhone in 2007 that is the first smartphone. It was integration of an audio player and a conventional mobile phone handset with a touch-screen based intuitive user interface. A smartphone is a commodity consumer product today and is a most promising personal information device in the future. Its display and the keypad are extracted and combined with a wrist watch to make a smart watch. "i'm" released the first smart watch, the "i'm Watch," in 2011. Another derivative from smartphones is a tablet PC represented by the iPad released in 2010. It is an upscale iPhone with a larger display. Evolution of personal information devices still continues with more applications to come.

VII. CONCLUSION

The origin of a variety of personal information devices, the Silicon Audio, has been presented with its family. It has

a memory card to store the compressed signal by MPEG-1/Audio Layer II algorithm. It has been shown that matured LSI technology to implement a decoder or even a codec on a single chip was essential to the Silicon Audio family in addition to the crossing of ever-increasing memory capacity and ever-decreasing bitrate necessary for satisfactory AV signal quality. With no mechanical movement, it is robust against shakes and vibrations. It has also changed the manufacturing process and companies therein. It is not simply the ancestor of the iPod, but has developed to a smartphone, a tablet PC, and a smart watch. Its evolution as a personal information device still continues with more applications to come.

REFERENCES

[1] "32 kbit/s adaptive differential pulse code modulation (ADPCM)," Rec. G.721, ITU-T, Nov. 1984.

[2] "7 kHz audio-coding within 64 kbit/s," Rec. G.722, ITU-T, Nov. 1988.

[3] "Extensions of Recommendation G.721 adaptive differential pulse code modulation to 24 and 40 kbit/s for digital circuit multiplication equipment application," Rec. G.723, ITU-T, Nov. 1988.

[4] "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)," Rec. G.726, ITU-T, Dec. 1990.

[5] "Coding of speech at 16 kbit/s using low-delay code excited linear prediction," Rec. G.728, ITU-T, Sep. 1992.

[6] G. A. Davidson, L. Fielder and M. Antill, "Low-Complexity Transform Coder for Satellite Link Applications," 89th Conv. of the Audio Eng. Society (AES), 2966, Sept. 1990.

[7] K. Tsutsui, H. Suzuki, O. Shimoyoshi, M. Sonohara, K. Akagiri and R. M. Heddle, "ATRAC: Adaptive Transform Acoustic Coding for MiniDisc," 93rd Conv. of the AES, 3456, Oct. 1992.

[8] G. C. P. Lokhoff, "dcc - Digital Compact Cassette," IEEE Trans. CE-37, No.3, pp.702–706, Aug. 1991.

[9] ISO/IEC 11172-3, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mb/s," Part 3: Audio, Aug. 1993.

[10] "Announcing Silicon Audio, Pocket-sized music player using semiconductor memory cards," NEC Press Release, December 1st, 1994. [on line] Available at http://www.nec.co.jp/press/en/9412/0101.html (accessed September 29th, 2017).

[11] "Announcing Silicon View, a hand-held video player with a semiconductor memory card," NEC Press Release, October 12th, 1995. [on line] Available at http://www.nec.co.jp/press/en/9510/1202.html (accessed September 29th, 2017).

[12] iPod (5th generation) in "Identify your iPod model," [on line] Available at https://support.apple.com/en-us/HT204217 (accessed September 29th, 2017).

[13] A. Sugiyama, M. Iwadare, N. Ohdate, T. Manabe, H. Takano, O. Kitabatake, E. Hirao, "The Silicon Audio, An audio-data compression and storage system with a semiconductor memory card," IEEE Trans. Consumer Elec., Vol. 41, No. 1, pp. 186–194, Feb. 1995.

[14] A. Sugiyama, M. Iwadare, T. Manabe, N. Ohdate, H. Takano, O. Kitabatake, E. Hirao, "A new implementation of the Silicon

Audio player based on an MPEG/Audio decoder LSI," IEEE Trans. Consumer Elec., Vol. 43, No. 2, pp. 207–215, May 1997.

[15] M. Iwadare, C. Kushiyama, and N. Ohdate, "A card audio player with trick-play functions," Proc. IEICE General Conf., A-4-58, pp. 167, Mar. 1998.

[16] S. Bergman et al., "The SR Report on the MPEG/Audio subjective listening test," ISO / IEC JTC1 / SC29 / WG11 MPEG91 / 010, May 1991.

[17] H. Fuchs, "Report on the MPEG/Audio subjective listening tests in Hannover," ISO / IEC JTC1 / SC29 / WG11 MPEG91 / 0331, Nov 1991.

[18] μ PD77220 user's manual, IEM-5078, Jun. 1988.

[19] M. Iwadare, H. Takano, Y. Shibuya, H. Sakamoto, T. Kuwajima, O. Kitabatake, N. Kobayashi, "A single-chip MPEG/audio decoder LSI based on a compact decoding algorithm," Proc. of IEEE Workshop on VLSI Sig. Proc., pp. 118–125, Sep. 1995.

[20] M. Iwadare, H. Takano, Y. Shibuya, H. Sakamoto, T. Kuwajima, O. Kitabatake, N. Kobayashi, "A single-chip MPEG/audio decoder LSI based on a compact decoding algorithm," J. of VLSI Sig. Proc., Vol. 16, pp. 25–30, 1997.

[21] ISO/IEC 13818-3, "ISO/IEC 11172-3 compatible low bit rate multi-channel coding system, and conventional stereo coding at lower sampling frequencies," Mar. 1995.

[22] P. P. Vaidyanathan, "Multirate systems and filter banks," Prentice-Hall, 1993.

[23] M. J. Narasimha and A. M. Peterson, "On the computation of the discrete cosine transform," IEEE Trans. COM-26, pp.934–946, Jun. 1978.

[24] M. Iwadare and T. Nishitani, "A fast subband analysis/synthesis algorithm for MPEG/Audio," 1992 IEICE Spring Nat'l Conv. Rec., A-194, Mar. 1992. (in Japanese)

[25] O. Kobayashi and H. Eda, "Flash EEPROM: Ready for use in external memory," Nikkei Electronics, pp.75–91, Apr. 11, 1994. (in Japanese)

[26] A. Sugiyama, "MPEG/Audio standardization and its ultimate application," Electronics, vol. 40, No. 6, pp. 56–59, Jun. 1995. (in Japanese)

[27] E. Grochowski and R. E. Fontana, Jr., "An analysis of flash and HDD technology trends," Flash Memory Summit, Santa Clara, CA, Aug. 2011.

[28] H. Matsuo, N. Ohdate, K. Mitsuhashi, and M. Iwadare, "Development of a Silicon View player based on an MPEG-1 audio/video decoder LSI," IEICE General Conf., A-198, Mar. 1996. (in Japanese)

[29] M. Iwadare, N. Ohdate, and H. Matsuo, "The portable Silicon View player," Proc. IIEEJ Ann. Conf., pp. 32–33, Jun. 1996.

[30] ISO/IEC 11172-2, "Coding of moving pictures and associated audio for digital storage media at up to about 1.5 Mb/s," Part 2: Video, Aug. 1993.

[31] Y. Katayama, I. Tamitani, A. Taniguchi, and Y. Ooi, "A single-chip MPEG1 Audio/Video decoder using macrocore and cell-base implementation," Proc. of IEEE Workshop on VLSI Sig. Proc., pp. 431–440, Oct. 1995.

[32] M. Iwadare, E. Sasaki, and N. Ohdate, "A portable multimedia guide terminal with a semiconductor memory card," Proc. IEICE Soc. Conf., A-4-30, p. 100, Sep. 1997.

[33] M. Iwadare, H. Matsuo, N. Ohdate, and M. Sato, "A shopping navigation terminal with user interactive functions," Proc. IEICE Soc. Conf., A-107, Sep. 1996.

[34] "Enjoy hifi sound while jogging," The Asahi, p. 3, Dec. 2, 1994.

[35] "Shock-proof music for 24 minutes with a business-card size semiconductor memory," The Nikkei, p. 12, Dec. 2, 1994.

[36] V. Houlder, "Music on the cards," Financial Times UK, p. 12, Dec. 2, 1994.

[37] "NEC unveils music by credit card," International Herald Tribune, p. 19, Dec. 2, 1994.

[38] "Technology you can use at home," Time, pp. 46–48, July 17, 1995.

[39] "Goodbye CDs?" Future Music, pp. 39–40, March 1995.

[40] ISO/IEC 14496-3, "Coding of audio-visual objects," Dec. 1999.

Biographies

Akihiko Sugiyama has been engaged in various research projects in signal processing such as audio coding and interference/noise control. He served as Chair of Audio and Acoustic Signal Processing Technical Committee, IEEE Signal Processing Society (SPS) [2011-2012], as associate editors for journals such as IEEE Trans. Signal Proc. [1994-1996], as the Secretary and a Member at Large to the Conference Board of SPS [2010-2011], as a member of the Awards Board of SPS [2015-2017], and as the Chair of Japan Chapter of SPS [2010-2011]. He was a Technical Program Chair for ICASSP2012. He has contributed to 16 chapters of books and is the inventor of 203 registered patents. He received 13 awards such as the 2006 IEICE Achievement Award, and the 2013 Ichimura Industry Award. He is Fellow of IEEE and IEICE, and a Distinguished Lecturer for IEEE SPS [2014-2015] and for IEEE CE (Consumer Electronics Society) [2017 -2018].

Masahiro Iwadare received the B.S. and the M.S. degrees in electrical engineering from the University of Tokyo, Tokyo, Japan, in 1984 and 1986, respectively. He joined NEC Corporation, Kawasaki, Japan, in 1986・詣 here he had been engaged in research and development as well as international standardization of speech and audio coding. Currently, he is a Manager in the Intellectual Property Management Division in charge of technology licensing. From October 1993 until September 1994, he was on leave at University of California, Santa Barbara, as a Visiting Researcher. His research interests include speech and audio signal processing. He was awarded the 1992 Shinohara Memoria1 Award by IEICE, the 2011 Promotion Foundation for Electrical Science and Engineering Award (the Ohm Award), the 2013 Ichimura Industry Award and the 2014 Prize for Science and Technology, the Commendation for Science and Technology by the Minister of Education, Culture, Sports, Science and Technology.
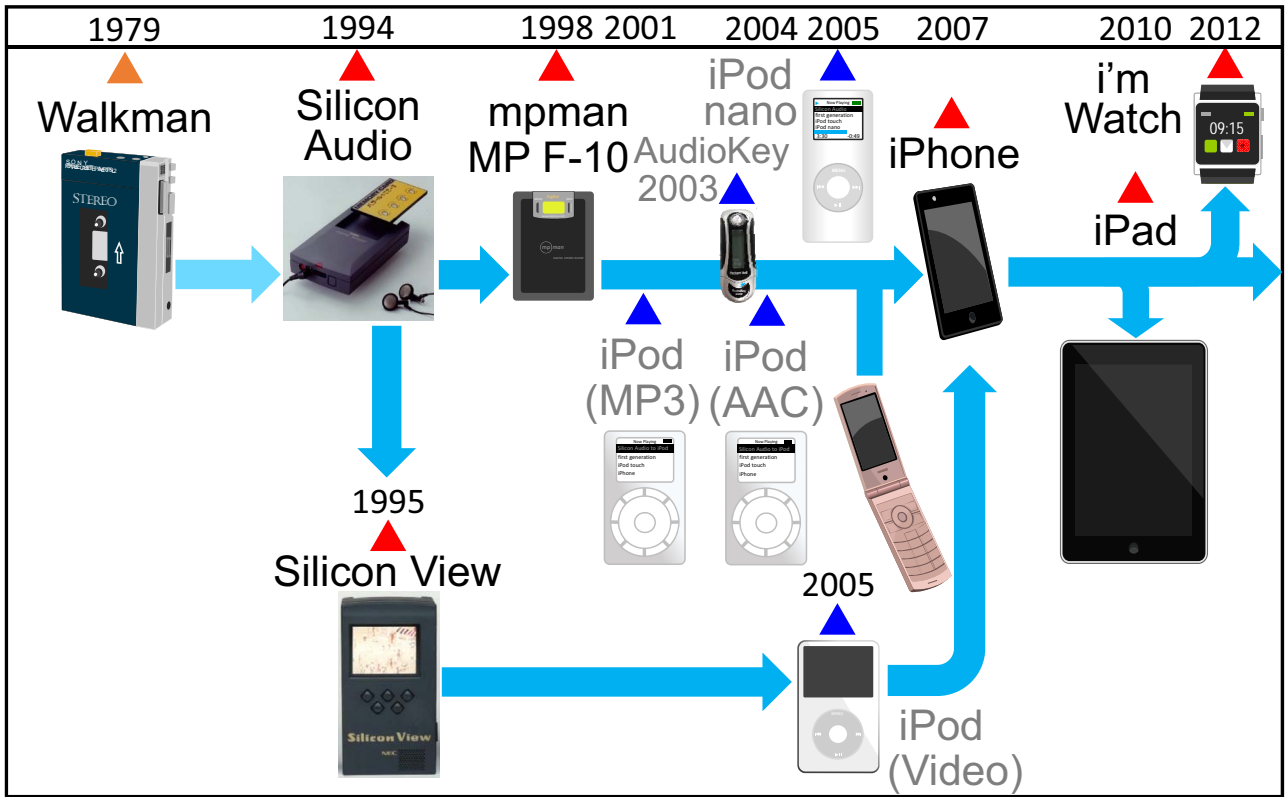
Fig. 18. A history of digital information devices: From the Walkman to the Silicon Audio, a smartphone, a tablet, and a smart watch.