5 Multimedia

In modern society, information processing systems are being used in various fields. Among these, the system that has frequently been used recently is the multimedia system. The multimedia system is a system that uses multimedia data such as moving images, still images, sound, and text by integrating it on a computer and can be represented in various forms.

5 - 1 Multimedia Technology

5-1-1 Multimedia

Multimedia refers to interactive (i.e., bidirectional) media. Information, such as characters, images, pictures, and sound, is digitalized (i.e., encoded) to integrate information media.
Different types of information available on the Internet that we use are registered on a Web server. (This is known as web content.) A large part of the web content is multimedia content in which characters, images, and sound are combined together. The software used for saving and managing the web content, and constructing a website is called CMS (Contents Management System). A typical CMS is Wiki by which the editing of a web page is performed easily from the browser. Wikipedia, which is called the encyclopedia of the Internet, is a typical usage example of Wiki.
Links (i.e., information for accessing related pages) are embedded in a large part of the web content, and different types of information can be searched for by following the links. Web content that is correlated through a link including text information is called hypertext, and web content (i.e., multimedia content) that is correlated through a link including images is called hypermedia.
In the authoring environment for creating the multimedia content, the software called multimedia authoring tool is mostly used. This software enables easy integration of material (i.e., files) such as documents, image/video, and sound through the mouse operation. At this time, PDF (Portable Document Format) may be used as one of the formats of document files. PDF is a format of document files developed by Adobe Systems that enables print images to be displayed in almost the same manner even in different applications by converting them into PDF files as they are, without any changes. When a print image is converted to a PDF file, the file size becomes smaller than the original data file, but editing cannot be performed in the original application. However, the characteristics of a PDF file also result in advantages such as prevention of data change and secondary use.
Because of its property, the data files of multimedia have a larger size than the data files of only text information. Therefore, on the Internet, the technique called streaming is used for

playing back at the download of a data file of video and music.

5-1-2 Sound Processing

Sound processing is a process that analog sound is handled as digital data. The techniques of A/D conversion by which an analog signal is encoded into a digital signal include PCM (Pulse Code Modulation). The procedure of encoding (i.e., A/D conversion) by PCM is as shown below.

[Encoding procedure by PCM]

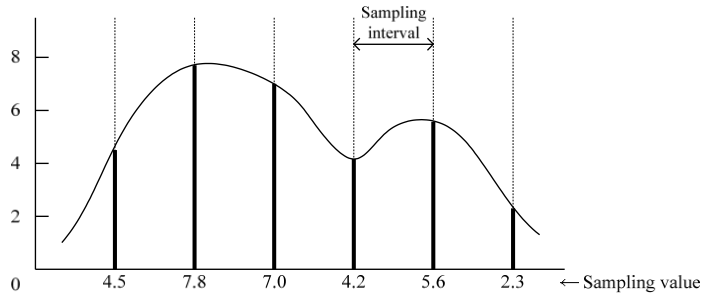
1) Sampling

The analog signals to be encoded are sampled at regular intervals. At this time, the sampling frequency indicating the sampling interval is determined on the basis of Shannon's sampling theorem.

- Shannon's sampling theorem:

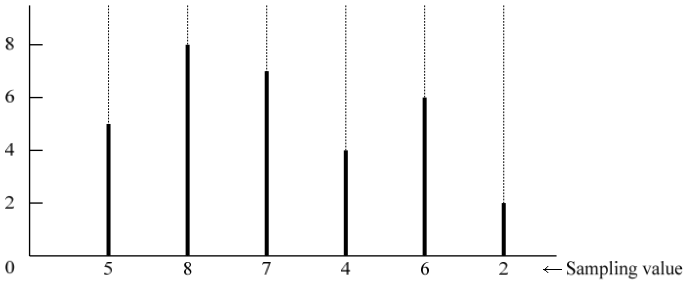
  When the highest frequency of the target analog signal is *f*, the original analog signal can be recovered if sampling is performed at a frequency of 2*f* or above.

  Note: Frequency is the number of vibrations in 1 second.
  The unit is Hertz (Hz).



2) Quantization

The sampled sampling values are rounded to the nearest integer numbers.



3) Encoding

A quantized integer number is encoded by representing it as a binary number. (The number of bits used as a code is called the quantization bit rate.)

Quantized value	5	8	7	4	6	2
	↓	↓	↓	↓	↓	↓
Encoded value	0101	1000	0111	0100	0110	0010

Example: When sound data that has the highest frequency of 2,000 Hz is encoded by a PCM method of 8-bit quantization bit rate, how many bits of digital data is contained in the entire sound data encoded in 1 second? Here, the sampling frequency is the minimum required frequency determined by Shannon's sampling theorem.

(1) Sampling frequency
= Highest frequency of analog signal (i.e., sound data) × 2
= 2,000 Hz × 2
= 4,000 Hz ··· Sampling is performed 4,000 times in 1 second.

(2) Number of bits obtained by encoding the sound data of 1 second
= Sampling frequency of 1 second × Quantization bit rate
= 4,000 times/second × 8 bits/one time
= 32,000 bits/second

As clear from the above example, the amount of digital data after encoding is determined on the basis of the sampling frequency (i.e., sampling interval) and quantization bit rate.
Moreover, D/A conversion that a digital signal (i.e., code) is restored as an analog signal (i.e., sound data) can be implemented by performing the encoding procedure in the reverse manner. At this time, in order to be able to perform restoration at a quality level that is closest to the original analog signal, either the sampling frequency is increased (sampling interval is shortened) and the number of sampling values is increased, or the quantization bit rate is increased to perform detailed staging of sampling values.

The typical formats (i.e., standards) of a sound file are as follows:

- MIDI (Musical Instruments Digital Interface)
  It is a file format that an electronic musical instrument and PC are connected and music data is exchanged. Not sound as such, but the rendition information for playing back the sound is handled.
- MP3 (MPEG1 Audio Layer3)
  It is a high-quality sound compression and decompression format using the sound

technology of MPEG (moving images compression and decompression format) standardized by ISO. It is also used in Internet music distribution and portable players.

- WAV (RIFF Waveform Audio Format)

  It is a sound data format that is mainly used in Windows systems. Various formats of data, such as ADPCM and WMA, can be stored in the container format.

  - ADPCM (Adaptive Differential Pulse Code Modulation):

    It is a sound signal compression method in which the PCM technique is applied.

  - WMA (Windows Media Audio):

    It is a streaming technique of music distribution that was originally developed by Microsoft. Recently, it has been used as a sound compression technique.

5-1-3 Still Image Processing

Still image processing is the process of handling images that do not move. The basic mechanisms of image representation include the color representation method and image quality.

- Color representation method

  This method uses the three primary colors of light (RGB) and the three primary colors of color (CMY). Generally, RGB is used in the color representation of the display, and CMY is used in the color representation of the printer. However, it is common to use the four colors of CMYK by adding black (Key) to the three primary colors of color.

- Image quality

  This is represented by resolution and gradation. Resolution represents the density of the pixels (i.e., dots) configuring an image in the number (dpi (dots per inch)) of pixels (i.e., dots). Gradation represents the stage of the color that can be represented in each pixel.

The typical formats (i.e., standards) of a still image file are as follows:

- BMP (Bit MaP)

  It is a format that represents the characters and images as a set of dots. A flaw that exists is that the amount of information generally increases more than in other formats.

  - TIFF (Tagged Image File Format)

    It is a format that enables handling of bit map images of various formats through the addition of tags (i.e., identifiers).

- JPEG (Joint Photographic Experts Group)

It is an international standard that supports full color and is defined by JPEG, an institution of ISO. It is a still image compression and decompression format that enables adjustment of the image quality and data amount by regulating the amount of compression. Because of an extremely high compression rate, it is used more frequently than other still image file formats. The compression formats of JPEG include lossless compression in which the original data is completely recovered and lossy compression in which the original data is not completely recovered. However, the lossy compression format is generally used.

- GIF (Graphic Interchange Format)

  It is a still image compression and decompression format that supports 256 colors. Although it is a lossless compression format, the amount of data is comparatively less, and it is also used in image processing on the WWW.

- PNG (Portable Network Graphics)

  It is a still image compression and decompression format that supports full color and is an expansion of GIF. It is a lossless compression format, and the amount of data is more than JPEG.

- Exif (Exchangeable Image File Format)

  It is a format that additional information (i.e., meta data), such as the photography time, is attached and saved in the image data taken by a digital camera.

5-1-4 Moving Image Processing

Moving image processing is the process of handling images that move. During the moving image processing of the computer, a series of still images is displayed in sequence. Each image is handled as a frame. At this time, the number of frames displayed in 1 second is the frame rate, which is represented in fps (frames per second).

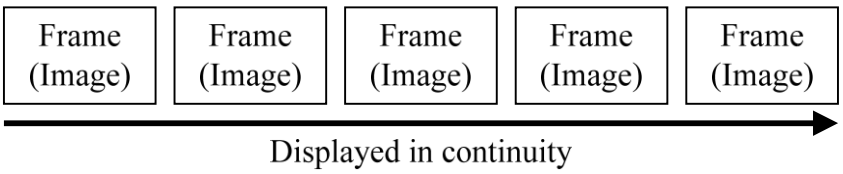


Figure 2-14 Image of moving image processing

The typical formats (i.e., standards) of a moving image file are as follows:

- MPEG (Moving Picture Experts Group)

  It is an international standard of moving image data that is defined by MPEG, an institution of ISO. It is a moving image compression and decompression format for which several standards are defined according to the quality of the image to be compressed.

MPEG-1: A compression and decompression technique of images that have a quality of video level

MPEG-2: A compression and decompression technique of television pictures and high-vision pictures

MPEG-4: A compression and decompression technique assuming the use of cellular communication

MPEG-7: A notational system of meta data assuming the implementation of a high-speed search engine

- QuickTime

  It is a moving image compression and decompression format used in Apple's "QuickTime." It makes use of the motion JPEG) method, where JPEG images are played back continuously, and can also be used in Windows.

- VRML (Virtual Reality Modeling Language)

  It is a file format for displaying 3D graphic data as pictures.

- AVI (Audio Video Interleaving)

  It is a moving image data format that is mainly used in Windows systems. It uses RIFF (Resource Interchange File Format), which is a format of multimedia files.

[CSV (Comma Separated Value) format]

CSV format is a file format used in spreadsheet software and database software. It is a format in which the data of each field is separated by a comma, and each record is separated by a linefeed code.

5-1-5 Compression and Decompression of Information

Because of their nature, the data files of multimedia have a larger size (amount of data) than the data files of only text information. If such a large data file is used as it is, a large recording area is required for storing the data file, and the load also increases when the file is exchanged across the network. Thus, in order to improve the efficiency of data storage and reduce the network load, the size (i.e., amount of data) of information is reduced in accordance with fixed rules (i.e., procedures). The process of reducing the size of information is called compression and the process of restoring the compressed information is called decompression.

The concept (i.e., format) of compression is broadly classified into the following two types of compression:

- Lossless compression

  It is a compression technique that enables a complete return to the original data from the compressed data.
- Lossy compression

  It is a compression technique that does not enable a complete return to the original data from the compressed data. In the case of images, since humans hardly ever notice the differences even when the restoration process is more or less incomplete, it does not result in a big problem. In comparison with lossless compression, the compression rate, which indicates the compression effect of the amount of data, can be increased; that is, the amount of data can be reduced.

The typical information compression and decompression formats (i.e., standards) are as follows:

- JPEG (Joint Photographic Experts Group)

  It is an international standard of a still image compression and decompression format that supports full color.
- MPEG (Moving Picture Experts Group)

  It is an international standard of a moving image compression and decompression format.
- ZIP

  It is a file compression format supported by several types of free software in addition to software products such as WinZip. It is used not only for compressing a single file but also for archiving, where several files are collectively compressed into one file. At an international level, it is a de facto standard (i.e., industry standard) file compression format.
- GZIP (GNU ZIP)

  It is a file compression format that is mainly used in UNIX. Although its name includes ZIP, it is not compatible with ZIP, and there is no archive function either.
- LZH

  It is a file compression format supported by the free software LHA. It is mainly used as a file compression format within Japan.
- MR (Modified Read) / MMR (Modified Modified Read)

  It is a compression and decompression format (i.e., encoding format) that is mainly used for facsimile.

5 - 2 Multimedia Application

(1) CG (Computer Graphics)

CG (Computer Graphics) is a technique of drawing images (i.e., graphics) and performing editing by using a computer.

[Software for CG]

- Painting software

  It is graphics software by which images are drawn in units of dots. The images thus created are called luster graphics. Since images are created by drawing lines with dots on the path traced by the mouse, this software can be used easily even by beginners.

- Drawing software

  It is graphics software by which images are drawn by calculating the direction and length (i.e., vector data) of lines. The images thus created are called vector graphics. The notches (jaggies) of the lines are not noticeable even when the drawn image is expanded.

[CG techniques]

- Anti-aliasing

  It is a technique by which the jaggies that occur in slanting lines and curved lines are made inconspicuous.

- Texture mapping

  It is a technique by which a texture is presented by pasting an image or a pattern onto the surface of a modeled object.

- Blending

  It is a technique by which the information on the degree of transparency is superimposed to represent a translucent image.

- Ray-tracing

  It is a technique by which a ray of light that has reached the eye point is traced in the reverse manner to perform rendering.

- Clipping

  It is a technique by which only the portion to be displayed is cut out from the entire image.

- Shading

  It is a technique of forming shadows on the surface of an object in order to provide the appearance of solidity.

- Morphing

It is a technique of changing the nodal points to smoothly change from one shape to another.

- Rendering

  It is a technique of displaying the data of an object as a two-dimensional image or picture.

- Polygon

  It is a polygon that is used to approximate a closed solid or a curved surface through three-dimensional CG.

- Key frame method

  It is a method for generating animation by interpolation between key frames.

(2) Three-dimensional picture (3D)

A three-dimensional picture (3D) is a picture (and the technique of creating pictures) that has depth and the appearance of solidity. It is created by three-dimensional computer graphics (3DCG) and motion capture, which uses a technique of attaching a sensor to each joint of a human body, feeding the information, such as coordinates and acceleration, detected by the sensors during movement to a computer, and then converting it into digital data. Moreover, in combination with virtual surround (e.g., a technique of virtually playing back 5.1-ch sound through two speakers or regular headphones), realistic content can be created.

(3) VR (Virtual Reality) / AR (Augmented Reality)

VR (Virtual Reality) refers to experiencing the virtual world that is generated by using CG as an actually existing world. On the other hand, AR (Augmented Reality) refers to extension of the real world by combining the real world and VR. In AR, information is added to actually existing things by VR and also emphasized.

(4) Other examples of application

A multimedia system that uses multimedia data, such as moving images, still images, sound, and text, by integrating it on a computer has several application examples.

- Internet broadcasting

  It refers to broadcasting through the medium of the Internet.
- Non-linear image editing system

  It is a method of editing images as digital data on a computer.
- CAD (Computer Aided Design)

  It is a computer system that supports the design of a product by using the CG techniques.
- Simulator

  It is a program or mechanism for a simulation approach to complex scientific phenomena that occur in reality. The prediction of the course of a typhoon and the wind tunnel experiment determining how the surrounding air flows when a car is running can be reproduced by the computer.
- Video game

  It is a game that can be played on the computer.
- Video on demand

  It is a service that distributes video according to the request of the user.
- Virtual mall (Cyber mall)

  It is a virtual shopping mall constructed on the Internet. Although there is no shopping mall, it is possible to actually buy real products.

Chapter 2 Exercises

Q1

In a client/server system, which of the following is the most appropriate function to be processed at the server side?

a) Process of displaying the output data
b) Process of updating the database
c) Checking the format of entered data
d) Process of displaying the pull-down menu

Q2

Which of the following is the computer system where one computer is in the standby state while the other computer operates normally?

a) Dual system
b) Duplex system
c) Multiprocessor system
d) Load sharing system

Q3

Which of the following is a component of a fault tolerant system?

a) RAID0
b) Duplexing of hard disk
c) Schedule backup
d) Data encryption

Q4

Which of the following is an appropriate description of the performance evaluation of a system?

a) In OLTP (On-Line Transaction Processing), the MIPS value is used in the performance evaluation of the system.
b) The response time and turnaround time are performance indicators from the viewpoint of the system operations manager.
c) If the utilization rate of system resources becomes high, the response time also

generally improves accordingly.

d) The number of transactions or jobs that can be processed within a unit time is important for evaluation of the system performance.

Q5

In a processor that has a clock time of 3 nanoseconds, when the number of clocks that is necessary for the execution of an instruction and the occurrence rate of the instruction are shown in the table below, what is the approximate performance (in MIPS) of this processor?

Type of instruction	Number of clocks necessary for the execution of an instruction	Occurrence rate
Register-to-register operation	4	40%
Register to/from memory operation	8	50%
Unconditional branch	10	10%

a) 5 b) 30 c) 50 d) 100

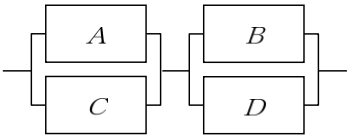
Q6

RASIS is a set of indexes or indicators for operating a system in a secure and stable manner. Among the indexes or indicators of RASIS, which of the following is the index that uses the probability of operation of the system?

a) Availability b) Integrity
c) Reliability d) Security

Q7

In the system configuration as shown in the figure below, which of the following is the closest availability of the overall system? Here, *A*, *B*, *C*, and *D* indicate the devices, and the availability is 0.9 for *A* and *C* and 0.8 for *B* and *D*.



a) 0.72 b) 0.92 c) 0.93 d) 0.95

Q8

Which of the following is the GUI tool that is used when only one item is selected from the multiple items that are mutually exclusive?

a) Scroll bar
b) Check box
c) Progress bar
d) Radio button

Q9

Which of the following is an appropriate purpose of appending a check digit to a customer code?

a) In order to detect an input error of the customer code
b) In order to arrange customers in order of their acquisition when a customer list is created
c) In order to enable classification of customers in groups on the basis of areas
d) In order to enable analogical inference of a specific customer

Q10

When a browser-based service of employee information is started, it is considered to publish a full-color face photograph of employees. Which of the following is the most appropriate image compression format for reducing the load on the in-house network?

a) GIF b) JPEG c) MIDI d) MPEG

Q11

Which of the following is the most appropriate explanation of virtual reality?

a) It refers to representing a world that is created within a computer as if it were a real world, by using CG techniques.
b) It refers to displaying an image in mosaic form first and then gradually displaying clearer for the purpose of improving the GUI, instead of sequentially displaying an image from the top.
c) It refers to a simulation approach to the wind tunnel experiment that is used for designing a car or an airplane through the use of a computer, and for testing whether or not the expected results are obtained.
d) It refers to supporting the design of a product on a computer by using CG techniques.