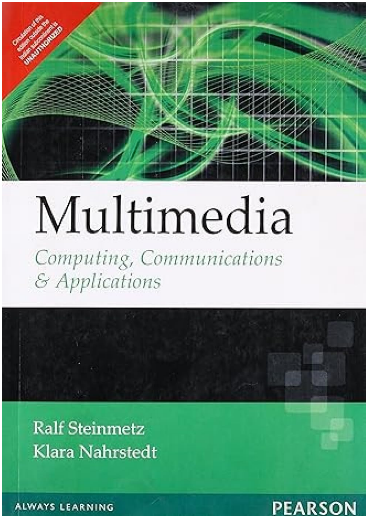
Multimedia Systems
Introduction to Multimedia

Er. Narayan Sapkota, M.Sc.
narayan.sapkota@eemc.edu.np

Everest Engineering College (Senior Lecturer)
Kathmandu University (Visiting Faculty)

November 18, 2024

Text Book


Multimedia
Computing, Communications
& Applications
Ralf Steinmetz
Klara Nahrstedt
ALWAYS LEARNING
PEARSON

Syllabus

1. Introduction
   - Overview of Multimedia
   - Global Structure of Multimedia*
   - Multimedia Building Blocks
   - Digital Representation
   - Interaction Techniques and Devices
2. The Medium Aspect
3. Main Properties of Multimedia System
4. Traditional Data Stream Characteristics
   - Data Stream Characteristics for Continuous Media*
5. Information Units

Table of Contents

Table of Contents

Overview of Multimedia (1)

The word multimedia is made up of two Latin-derived components: the prefix multi- and the root media. The prefix multi- comes from the Latin word multus, which means "numerous" or "many." The root media, however, has a more complex background. Media is the plural form of the Latin word medium, meaning "middle" or "center." Over time, medium evolved to describe a channel or means of communication, such as radio, television, or print. In this context, media refers to the various forms through which information, ideas, or art can be shared.

Multimedia literally translates to "multiple means or channels of communication," describing any system that combines text, audio, video, and other formats to convey information. Therefore, multimedia is a mixture of different forms of media: text, graphics, audio, video, animation.

Overview of Multimedia (2)

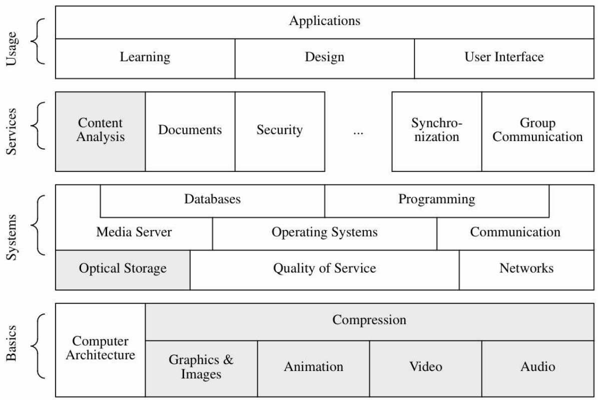


Figure 1: The most important multimedia fields

Overview of Multimedia (3)

1. **Basics:** In addition to the computer architecture for multimedia systems, one of the most important aspects is a media-specific consideration.
2. **Systems:** This section covers system aspects relating to processing, storage, and communication and the relevant interfaces.
3. **Services:** This section details single functions, which are normally implemented through individual system components.
4. **Usage:** This section studies the type and design of applications and the interface between users and computer systems.

Overview of Multimedia (4)

Information in multimedia formats is often stored on CDs and DVDs, making it accessible and easily distributable. Fundamental feature of multimedia is interactivity–the ability of the user to interact with an application. Examples include keyboard and mouse input, mouse rollovers, voice activation, and touch screen.

Multimedia means, from the user's perspective, that computer information can be represented through audio and/or video, in addition to text, image, graphics and animation.

For example, using audio and video, a variety of dynamic situations in different areas, such as sport or ornithology lexicon, can often be presented better than just using text and image alone.

Table of Contents

Global Structure of Multimedia* (1)

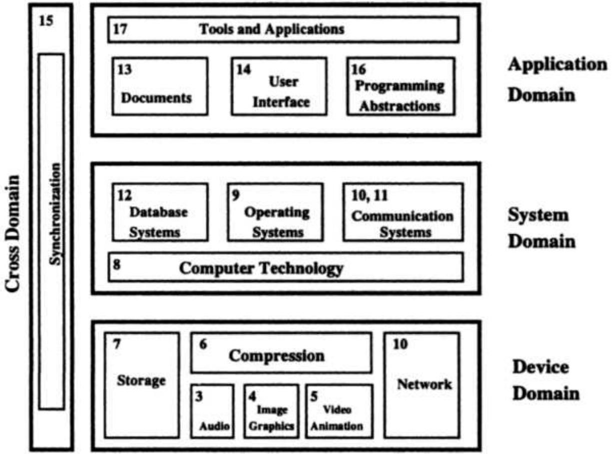


Figure 2: Global Structure of Multimedia

Global Structure of Multimedia* (2)

1. **Device Domain:** Basic concepts for processing of digital audio and video data are based on digital signal processing. Hence, these concepts are described and some possible practical implementations are presented. Different methods for the processing of image, graphics and animations are described. The audio techniques section includes music (MIDI) and speech processing. The understanding of video techniques is built mainly on TV development, including digital representation and HDTV. The originated data rates of these media demand, because of the current quality requirements and available technology, corresponding compression methods. The diminishing cost of optical storage space has contributed significantly to the current development of computer technology. Almost all developments are based on CD-DA(Compact Disc-Digital Audio). On the other-hand, networks, with their higher bandwidth and capacity for transmitting all media types, have led to networked multimedia systems.

Global Structure of Multimedia* (3)

2. **System Domain:** The interface between the device domain and the system domain is specified by the computer technology. To utilize the device domain, several system services are needed. Basically, three services exist. These services are mostly implemented in software:
   - The *operating system* serves as an interface between computer hardware/system software and all other software components. It provides the user with a programming and computational environment, which should be easy to operate. In its function as an interface, the operating system provides different services that relate to the computer resources, such as: processor, main memory, secondary storage, input and output devices and network.
   - The *database system* allows a structured access to data and a management of large databases.
   - The *communication system* is responsible for data transmission according to the timing and reliability requirements of the networked multimedia application.

Global Structure of Multimedia* (4)

3. **Application Domain:** The services of the system domain are offered to the application domain through programming abstractions. Moreover, such abstractions can be, for example, part of a multimedia operating system, programming language or object-oriented class hierarchy. Another topic embedded in the application domain is document handling. A document consists of a ere of structured information, represented in different media, and generated or recorded at the time of presentation. Many functions of document handling and other applications are accessible and presented to the veer through a user interface.
4. **Cross Domain:** It tune out that some amperes, such as synchronization aspects, are difficult to locate in one or two components or domains. The reason is that synchronization, being the temporal relationship among various media, relates to many components across all domains

Table of Contents

Multimedia Building Blocks (1)

Any multimedia application consists any or all of the following components:

1. **Text:** Text is a fundamental component of multimedia, acting as a primary vehicle for information and communication. In the context of the Web, text is crucial, with HTML enabling the display of text on computer screens. HTML was initially designed to showcase simple documents, but over time, it evolved to support interactive elements, layout designs, and multimedia integrations. Text can vary from headlines, body paragraphs, or even decorative fonts, and is essential in providing context, instructions, or narratives within a multimedia application.

Multimedia Building Blocks (2)

2. **Audio:** Audio in multimedia introduces a powerful auditory experience, enhancing user engagement through music, spoken words, sound effects, and ambient sounds. It can serve various purposes: setting a mood, emphasizing critical points, narrating information, or simply adding entertainment value. Whether it's a voice-over explaining content, a background score that sets the tone, or a sound effect that accentuates an action, audio provides a deeper layer of immersion and often helps retain user interest by stimulating the sense of hearing.

Multimedia Building Blocks (3)

3. **Images:** Images bring a visual richness to multimedia that can make complex ideas easier to understand and more appealing. Represented either in analog (traditional photographs) or digital formats (such as JPEG, PNG, or SVG), images include still photos, illustrations, and graphic designs. They can serve as visual representations of data, enhance textual information, or simply add aesthetic appeal. In digital multimedia, images can be edited, enhanced, or layered to produce visually compelling content, improving the overall communicative power of the application.

4. **Animation:** Animation is the rapid sequence of images, 2-D artwork, or 3-D models that creates the illusion of movement, leveraging the phenomenon known as "persistence of vision." In multimedia, animation is a versatile tool for demonstrating processes, visualizing concepts, and storytelling. It can range from simple GIF animations to sophisticated 3-D motion graphics and simulations. Animation enhances the user experience by illustrating actions that might be too complex or dynamic to convey through static images, thus adding life to the multimedia presentation.

Multimedia Building Blocks (5)

5. **Video:** Video combines audio, text, and visuals in a moving format that allows for the realistic portrayal of dynamic scenes, often recorded in real time.Digital video, the standard today, offers high quality and is easily integrated into various multimedia platforms.Video can serve educational, promotional, entertainment, or instructional purposes, adding realism and depth to the presentation. Video enables a richer storytelling format and provides real-time depictions of events, procedures, or demonstrations, making it an essential component of interactive multimedia.

Table of Contents

Digital Representation of Multimedia (1)

Digital representation of multimedia refers to encoding various multimedia components—such as text, audio, images, graphics, animation, and video—into binary code (0s and 1s) so that they can be processed, stored, and transmitted by digital devices.

Each type of media requires specific methods of encoding and compression to ensure quality, efficiency, and compatibility across different platforms.

1. **Text:** Text is represented digitally using character encoding standards like ASCII (American Standard Code for Information Interchange) or Unicode. ASCII uses 7 or 8 bits to represent each character, while Unicode, which supports a wider range of languages and symbols, can use up to 32 bits per character. This binary encoding allows text to be displayed, processed, and formatted across different digital platforms and devices.

Digital Representation of Multimedia (2)

2. **Audio:** Audio is represented digitally by sampling sound waves at regular intervals and converting these samples into binary data. Two key factors in audio digitization are:
    - *Sampling Rate:* The number of samples taken per second, measured in Hertz (Hz). Higher sampling rates (e.g., 44.1 kHz for CD quality) capture more detail.
    - *Bit Depth:* The number of bits used per sample, affecting the audio's dynamic range and fidelity. Common bit depths are 16-bit (CD quality) and 24-bit (studio quality).

    Common digital audio formats include MP3, WAV, and AAC, each with its own compression algorithms to balance quality and file size.
3. **Image:** Digital images are represented as a grid of pixels, where each pixel has a specific color value. There are two main types of digital images:

Digital Representation of Multimedia (3)

- *Raster Images:* Made up of pixels, commonly used in formats like JPEG, PNG, and BMP. Each pixel's color is represented in bits, often in RGB (Red, Green, Blue) format.
- *Vector Images:* Made up of paths defined by mathematical equations rather than pixels, often used in formats like SVG. Vector images are resolution-independent, making them ideal for logos and graphics that need to be scaled without losing quality.

4. **Graphics:** Graphics in multimedia may be vector-based (like SVG files) or raster-based (like PNG or JPEG). Vector graphics are resolution-independent, stored using mathematical formulas, and can be scaled without losing clarity. Raster graphics, by contrast, are stored as a grid of pixels and can lose quality when resized. Digital representation of graphics is essential for interactive applications, where images and illustrations need to adapt to different screen sizes and resolutions.

Digital Representation of Multimedia (4)

5. **Animation:** Digital animations are represented as a series of sequential frames. Each frame contains an image that, when shown in rapid succession, creates the illusion of motion. There are two primary types of digital animation:
   - *Frame-by-Frame Animation:* Each frame is an individual image (e.g., in GIFs).
   - *Keyframe Animation:* Only specific "key" frames are stored, with the software interpolating the frames in between.

   Animation formats like GIF, SWF, and HTML5 Canvas are common in multimedia, while advanced animations and 3D animations may use formats like MP4 or proprietary formats for platforms like Unity and Blender.

Digital Representation of Multimedia (5)

6. **Video:** Video is represented digitally as a sequence of images (frames) displayed at a specific frame rate, typically measured in frames per second (fps). Each frame is a complete image that, when played in sequence, provides the illusion of motion. Key factors in digital video representation include:
   - *Resolution:* The number of pixels in each frame, e.g., 1920x1080 for Full HD.
   - *Frame Rate:* The number of frames shown per second, e.g., 24 fps or 30 fps.
   - *Compression:* Video files are often large, so formats like MP4, AVI, and MOV use compression techniques (such as H.264 or H.265) to reduce file size while maintaining quality.

Benefits of Digital Representation

- *Storage and Retrieval:* Digital data can be stored compactly and retrieved quickly, making it accessible anytime on compatible devices.

Digital Representation of Multimedia (6)

- *Compression and Efficiency:* Compression algorithms enable multimedia data to occupy less space, making transmission faster and storage more efficient.
- *Interactivity and Manipulation:* Digital multimedia allows interactive elements, such as hyperlinks, animations, and dynamic graphics, which are harder to achieve with analog media.
- *Consistency and Quality:* Digital representations offer consistent quality over time, with little to no degradation due to copying or transmission.

Table of Contents

Interaction Techniques and Devices (1)

Interaction techniques and devices enable effective communication between users and multimedia content. These techniques and devices enhance user experience through responsiveness, immersion, and ease of control.

1. **Pointing Devices:** Pointing devices are fundamental for navigating and interacting with on-screen elements, particularly in GUIs. These devices allow users to select, drag, and drop objects, facilitating direct interaction with multimedia content.
   - *Mouse:* used for navigation, selection, and interaction via clicking and dragging.
   - *Trackpad/Touchpad:* Typically on laptops, supports multi-touch gestures like pinch and zoom.
   - *Stylus:* Precision tool for drawing, annotating, and detailed interaction on touchscreens and tablets.

Interaction Techniques and Devices (2)

2. **Touchscreen Interaction:** Touchscreens have revolutionized interaction by allowing direct, hands-on engagement with content. They support various gestures and taps for intuitive control and are widely used in mobile devices, tablets, kiosks, and interactive displays.
   - *Single-Touch:* Basic interaction through taps or swipes.
   - *Multi-Touch:* Enables advanced gestures like pinch, zoom, and rotation.
   - *Haptic Feedback:* Adds tactile feedback, enhancing the sense of touch during interactions.
3. **Voice Interaction:** Voice interaction allows users to control multimedia systems hands-free, using speech commands to operate devices or navigate content. It's particularly useful for smart devices, virtual assistants, and accessibility features.
   - *Voice Recognition:* Systems like Siri and Alexa recognize spoken commands for navigation and control.
   - *Speech-to-Text:* Converts speech to text for dictation and messaging.

Interaction Techniques and Devices (3)

- *Natural Language Processing (NLP):* Enables conversational and context-aware interaction.

4. **Gesture Recognition:**Gesture recognition interprets physical movements or gestures as commands, allowing users to interact with devices without direct contact. This method is often used in gaming, augmented reality (AR), virtual reality (VR), and smart home environments.
   - *Camera-Based Recognition:* Devices like Kinect detect body movements for interaction.
   - *Wearable Sensors:* Motion-sensing gloves enable precise control in VR and AR.
   - *Touchless Gestures:* Simple hand gestures (e.g., swipes) enable contactless interaction.
5. **Virtual Reality (VR) and Augmented Reality (AR) Devices:** R and AR devices provide immersive environments, transforming the way users experience multimedia by engaging their senses more directly.

Interaction Techniques and Devices (4)

- *VR Headsets:* Fully immersive 3D experience with tracked hand controllers for interaction.
- *AR Glasses:* Overlay digital content onto the real world, blending virtual and physical spaces.
- *Motion Controllers:* Allow physical actions for realistic VR and AR engagement.

6. **Keyboard Interaction:** The keyboard remains a primary device for text input and various shortcuts. Multimedia systems often support keyboard shortcuts to speed up navigation and control, enhancing productivity and accessibility.
   - *Standard Keyboard:* Primarily used for text input and navigation.
   - *Gaming Keyboards:* Equipped with programmable keys and optimized for precise controls.
   - *Virtual Keyboards:* On-screen keyboards for touch or voice-activated systems.

7. **Game Controllers and Joysticks:** Game controllers and joysticks provide specialized input for gaming and other interactive multimedia applications, often supporting haptic feedback and motion sensing for immersive engagement.
   - *Standard Game Controllers:* Buttons, triggers, and joysticks for precise interaction.
   - *Motion-Sensing Controllers:* Devices like the Wii Remote that respond to physical movement.
   - *Joysticks:* Often used in simulators for detailed control over motion and speed.
8. **Sensor-Based Interaction:** Sensors detect environmental data (such as touch, proximity, pressure, or temperature) and allow users to interact with devices through non-traditional means. They are commonly used in interactive installations, smart home systems, and mobile devices.

Interaction Techniques and Devices (6)

- *Proximity Sensors:* Detect nearby objects, enabling gesture-based interaction.
- *Accelerometers and Gyroscopes:* Track movement and orientation for dynamic control.
- *Environmental Sensors:* Used for real-time data in smart devices, such as temperature or pressure sensors.

Table of Contents

The Medium Aspect (1)

Medium is defined as means for distribution and presentation of information. Examples of a medium are text, graphics, speech, and music. Media can be classified with respect to different criteria. We classify media according to perception, representation, presentation, storage, transmission, and information exchange.

1. **Perception Media:** Perception media refers to the nature of information perceived by humans, which is not strictly identical to the sense that is stimulated.For example, a still image and a movie convey information of a different nature, though both stimulate the same sense. The key question is: How do humans perceive information?

   In this context, we primarily distinguish between what we see and what we hear. Auditory media include music, sound, and voice, while visual media include text, graphics, and still and moving pictures. This differentiation can be further refined. For example, a visual medium can consist of moving pictures, animation, and text. Moving pictures, in turn, are typically composed of a series of scenes, which are made up of individual pictures.

2. **Representation Media:** The term representation media refers to how information is represented internally to the computer, with the encoding method being crucial. The key question is: How is information encoded in the computer? There are several encoding options:
   - Each character of a text can be encoded in ASCII.
   - A picture can be encoded using standards like CEPT, CAPTAIN, or the GKS graphics standard.
   - An audio data stream may use simple PCM encoding with a linear quantization of 16 bits per sampling value.
   - A single image may be encoded in Group-3 facsimile format or JPEG format.

The Medium Aspect (3)

- A combined audio-video sequence can be stored using various TV standards (e.g., PAL, SECAM, or NTSC), the CCIR-601 standard, or in MPEG format.

3. **Presentation Media:** The term presentation media refers to the physical means used by systems to reproduce information for humans. For example, a TV set uses a cathode-ray tube and loudspeakers. The key question is: Which medium is used to output or input information from the computer?

   We primarily distinguish between output and input media. Media such as paper, computer monitors, and loudspeakers are considered output media, while keyboards, cameras, and microphones are input media.

The Medium Aspect (4)

4. **Storage Media:** The term storage media is often used in computing to refer to various physical means for storing computer data, such as magnetic tapes, magnetic disks, or digital optical disks. However, data storage is not limited to the components available in a computer, meaning that paper is also considered a storage medium. The key question is: Where is information stored?
5. **Transmission Media:** The term transmission media refers to the physical means—such as cables of various types, radio towers, satellites, or the ether (the medium that transmits radio waves)—that allow the transmission of telecommunication signals. The key question is: Which medium is used to transmit data?

The Medium Aspect (5)

6. **Information Exchange Media:** Information exchange media include all data media used to transport information, such as storage and transmission media. The key question is: Which data medium is used to exchange information between different locations?

   For example, information can be exchanged by storing it on a removable medium and transporting the medium from one location to another. These storage media include microfilms, paper, and floppy disks. Information can also be exchanged directly through the use of transmission media, such as coaxial cables, optical fibers, or radio waves.

7. **Representation Values and Representation Space:** The concept of "medium" in information processing relates to how information is perceived by the human senses. Perception media primarily address the five senses, defining presentation values within presentation spaces. Visual presentation spaces include paper or computer monitors, while acoustic spaces include stereophony and quadrophony. These presentation spaces are used to output information.

   Presentation values represent information differently depending on the medium. For example, text represents sentences as sequences of characters, while voice represents sound as pressure waves. Some media, like temperature or smell, cannot be directly interpreted by humans, while others, like text and voice, require predefined symbols to understand the information. Presentation values can be continuous, as in sound or light waves, or discrete, as in text characters and audio signal samples.

8. **Representation Dimensions:** Each representation space consists of one or more representation dimensions. A computer screen has two spatial dimensions; holography and sterophony require ad additional spatial domain. Time can occur inside each representation space as an additional dimension. Media is of two types with respect to time in their representation space: Time-Independent and Time-Dependent Media

The Medium Aspect (8)

1. *Time-Independent (or Discrete) Media:* In this type, individual elements do not depend on time to convey information. Their content consists of distinct elements (or values) that are independent of timing. Text, graphics, and static images are typical examples of time-independent media. For instance, a page of text in a book does not require a specific time-frame to be understood and can be processed as quickly or slowly as desired without affecting the information's validity.
   Discrete in Nature: These media can be viewed as a sequence of individual elements that do not change over time.

   Non-Critical Processing: Processing can happen as fast as possible without concern for timing, as the correctness of the data doesn't depend on when it's accessed or displayed.

   The term "discrete" can be confusing because some media may be discrete in value but still time-dependent in nature. For example, a series of video frames (each a discrete image) requires specific timing to be displayed continuously.

2. *Time-Dependent (or Continuous) Media:* The timing of individual elements is essential. Information is conveyed not only by the individual elements but also by their timing and sequence. Audio, video, and sensor data are typical examples. For instance, the meaning of an audio sample depends on when it occurs relative to other samples.

   Continuous Representation: These media require a continuous sequence of values over time, making the timing critical for accuracy.

   Time-Critical Processing: Data must be processed in a timely manner because delays can result in incorrect or lost information. For example, a delayed audio sample may lose meaning if subsequent samples have already played, disrupting the flow.

Table of Contents

Main Properties of Multimedia System (1)

According to [2] key properties of multimedia system are–Discrete and continuous Media, Independent Media, Computer Controlled Systems, and Integration. According to [1](our textbook) key properties are multimedia system definition, combination of media, independence, computer-supported Integration, and communication systems.

1. **Multimedia System Definition:** A multimedia system, based on the traditional definition, is any system that supports more than one type of media, such as text, graphics, sound, and video. However, this quantitative approach is insufficient because it only considers the number of media types, not the quality or nature of their integration.For example, a system that processes both text and graphics would be classified as multimedia, even though such systems existed before the concept of multimedia in computer environments.

Main Properties of Multimedia System (2)

Thus, the term "multimedia" implies a new quality in the context of computer systems, not just a greater variety of media. The focus should be on the type and integration of media rather than simply their quantity. This understanding of multimedia is more qualitative.However, this definition is not universally accepted, and standardization bodies like ISO often use a weaker interpretation of the term. Key distinguishing properties of a multimedia system include the combination of different media types, media-independence, computer control, and integration, which set multimedia systems apart from other systems.

2. **Combination of Media:** Not every arbitrary combination of media justifies the usage of the term multimedia. A simple text processing program with incorporated images is often called a multimedia application because two media are presented through one program. However, the term multimedia should only be used when both continuous and discrete media are utilized. Therefore, a text processing program with incorporated images is not considered a multimedia application.
3. **Independent Media:** An important aspect is that the media used in a multimedia system should be independent. Although a computer-controlled video recorder handles audio and moving image information, there is a temporal dependence between the audio part and the video part. In contrast, a system that combines signals recorded on a DAT (Digital Audio Tape) recorder with some text stored in a computer to create a presentation meets the independence criterion. Other examples

are combined text and graphics blocks, which can be in an arbitrary spatial arrangement in relation to one another.

4. **Computer Supported Integration:** The media-independence principle allows for the flexible combination of different media, with computers being the ideal tools for this purpose. A multimedia system should be capable of computer-controlled media processing, where the integration of independent media types can be programmed by system programmers or even users. Simple input or output of various media types, such as through a video recorder, does not satisfy the requirement for a truly integrated solution. Real integration involves synchronizing timing, spatial, and semantic relationships between media. High integration is demonstrated when, for example, changes in the content of a table can trigger corresponding changes in both video and text. Such a level of integration is rare, even in advanced multimedia systems. An integrated multimedia system allows for much more flexibility, enabling

functions like sending audio or combined audio and text messages, in addition to the traditional text-based communication of conventional systems. This emphasizes the importance of flexible media processing and high integration in multimedia environments.

5. **Communication Systems:** Communication-capable multimedia systems must be considered. A reason for this is that most of today's computers are interconnected, and considering multimedia mainly from a local processing viewpoint would be a limitation, if not a step backward. Another reason is that distributed environments enable particularly interesting multimedia applications. In such environments, multimedia information can not only be created, processed, presented, and stored, but also distributed beyond the boundaries of a single computer.

Table of Contents

Traditional Data Stream Characteristics [1] (1)

We defined multimedia from a computer-based perspective in above sections. However, this section includes the consideration of multimedia communication systems. Therefore, we need to specify the notion of multimedia from a communication point of view. In distributed multimedia communication, data from both discrete and continuous media are transmitted, with information exchange occurring between a source and destination, either on the same or different machines.

Data is divided into packets, forming time-dependent data streams, which can contain either continuous (e.g., speech) or discrete (e.g., document retrieval) media. Different types of media transmission result in distinct data stream features. The concepts of asynchronous, synchronous, and isochronous data transmission, originating from computer communication, are used in systems like Fiber Distributed Data Interface (FDDI) networks to describe data transmission modes based on end-to-end packet delay.

Traditional Data Stream Characteristics [1] (2)

1. **Asynchronous Transmission Mode:** In the broadest sense, communication is called asynchronous if the sender and receiver do not need to coordinate before data transmission begins. Data can be sent at any instant, with bit synchronization provided by two independent clocks, one at the sender and the other at the receiver.

   An example of asynchronous transmission is the way simple ASCII terminals connect to host computers. Each time the character "A" is pressed, a sequence of bits is generated at a preset speed. To notify the computer interface that a character is arriving, a special signal called the start signal– often not a bit itself– precedes the information bits. Similarly, a stop signal follows the last information bit.

Traditional Data Stream Characteristics [1] (3)

2. **Synchronous Transmission Mode:** The term "synchronous" refers to the relationship between two or more repetitive signals that have simultaneous significant instants. Data transmission can only begin at well-defined times that align with a clocking signal synchronized with the receiver's clock. Clocked transmission is crucial, especially for time-sensitive data.

   For example, consider a digitized voice signal transmitted across a non-synchronized network: as network traffic increases, the transmission may face delays, causing the data stream to slow down when traffic is heavy and speed up when it decreases. In a digitized phone call, if audio samples are delayed, the listener may experience this delay as disruptive interference or noise. Since the receiver cannot accelerate playback to catch up with delayed samples, synchronous transmission

Traditional Data Stream Characteristics [1] (4)

ensures smoother, more consistent data flow for time-dependent applications.

3. **Isochronous Transmission Mode:** The term "isochronous" refers to a periodic signal in which the time intervals between corresponding transitions are consistent, either equal to the unit interval or a multiple of it. Isochronous transmission ensures a constant phase relationship between the significant instants of two or more sequential signals. In this mode, individual characters are separated by whole-number bit-length intervals, unlike asynchronous transmission, where characters may have variable-length intervals between them.

   For example, an end-to-end network connection is considered isochronous if the bit rate is guaranteed and the delay jitter is both predictable and minimal. Isochronous transmission is essential for networks that need to support continuous media streams, such as real-time audio or video.

Traditional Data Stream Characteristics [1] (5)

To transport audio and video in real time, the network must sustain the bit rate required by the source, ensuring smooth and uninterrupted data flow.

Table of Contents

Data Stream Characteristics for Continuous Media* (1)

This section summarizes the characteristics of data streams in multimedia systems, focusing on audio and video transmissions. It describes the impact of compression methods applied to these data streams prior to transmission. This classification is relevant for both distributed and local environments.

1. **Strongly and Weakly Periodic Data Streams:** The first property of data streams concerns the time intervals between the completed transmissions of consecutive packets or information units. Based on when packets are ready, we can identify the following types:
   - *Strongly Periodic Data Stream:* In this type, the time interval between consecutive packets is constant, meaning there is minimal jitter—ideally zero. An example is PCM-encoded (Pulse Code Modulation) voice in telephone systems, as illustrated in Figure 3.

Data Stream Characteristics for Continuous Media* (2)

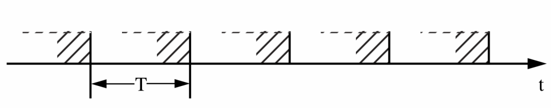


Figure 3: Strongly periodic data stream; time intervals have the same duration be tween consecutive packets

- *Weakly Periodic Data Stream:* The duration of time intervals between consecutive packets is often represented as a function with a finite period. However, unlike a strongly periodic data stream, this interval is not constant between neighboring packets. As shown in Figure 4.

Data Stream Characteristics for Continuous Media* (3)

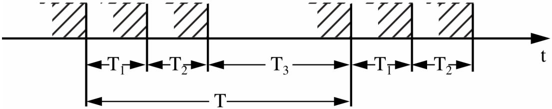


Figure 4: Weakly periodic data stream; time intervals between consecutive packets are periodic

- *Aperiodic Data Streams:* All other transmission types are referred to as aperiodic data streams, where the intervals between packets vary without a regular pattern, as illustrated in Figure 5. An example of an aperiodic data stream is a multimedia conferencing application with a shared screen window. In such applications, data like mouse position and button status (e.g., "left button pressed") are only transmitted when changes occur. Periodic transmission in this case would create excessive data redundancy and an unnecessarily high data rate. An ideal system transmits data only when there is an active change in position or status.

Data Stream Characteristics for Continuous Media* (4)

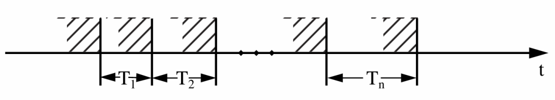


Figure 5: Aperiodic data stream; the time interval sequence is neither constant nor weakly periodic

2. **Variation of the Data Volume of Consecutive Information Units:** A second key characteristic of data streams is the variation in the quantity of data in consecutive packets or information units.
    1. *Strongly Regular Data Stream:* If the data volume remains constant throughout the data stream's lifetime, it is considered a strongly regular data stream, as shown in Figure 6. This is typical for uncompressed digital audio-video streams, such as a full-image encoded data stream from a camera or audio from a CD.

Data Stream Characteristics for Continuous Media* (5)

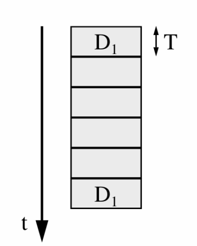


Figure 6: Strongly regular data stream; the data quantity is constant in all packets

Data Stream Characteristics for Continuous Media* (6)

2. *Weakly Regular Data Stream:* If the data volume changes periodically over time, it is classified as a weakly regular data stream (see Figure 7). Certain video compression methods, like MPEG, encode full images individually (I frames), creating large data packets. Between these I frames, smaller packets (P and B frames) carry only the differences between consecutive images, resulting in a fluctuating but generally periodic data volume. For MPEG, a typical average I:B data ratio is 10:1:2, producing a weakly regular data stream.

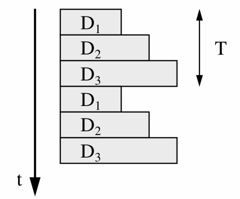


Figure 7: Weakly regular data stream; the packets' data stream varies periodically

Data Stream Characteristics for Continuous Media* (7)

3. *Irregular Data Stream:* When the data volume varies unpredictably and does not follow a periodic function (see Figure 8), the data stream is considered irregular. This type of data stream, often produced by variable bit rate compression, is more challenging to transmit and process. Here, the size of each packet depends on the specific changes in image content compared to the previous frame, creating significant variability in data size based on the video sequence content.

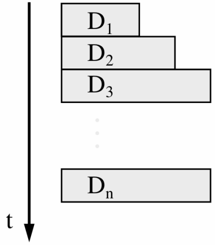


Figure 8: Irregular data stream; the packets' data quantity is not constant and does not vary periodically

3. **Interrelationship of Consecutive Packets:** The third characteristic for qualifying data streams concerns the continuity or relationship between consecutive packets. Are packets transmitted consecutively, or are there gaps between them? This can be described by examining how the resources, such as the network, are utilized.
   1. *Interrelated Data Stream:* As shown in Figure 9, an interrelated data stream involves consecutive packets transmitted one after another without any gaps. Additional or layer-independent information, such as error detection codes, is included, ensuring 100% utilization of the resource. This allows maximum throughput and optimal resource usage. An example is an ISDN B channel transmitting audio data at 64 Kbit/s.

Data Stream Characteristics for Continuous Media* (9)

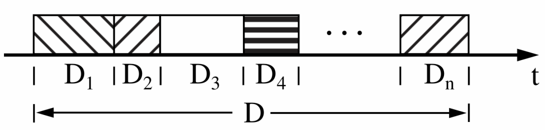


Figure 9: Interrelated data stream; packets are transmitted without gaps in between

2. *Non-interrelated Data Stream:* When packets are transmitted over a higher-capacity channel, gaps may occur between them, resulting in a non-interrelated data stream. As shown in Figure 10, these gaps can vary in duration, and it doesn't matter if they appear between all packets. An example of this is the transmission of a data stream encoded by the DVIPLV method over an FDDI network, where the average bit rate of 1.2 Mbit/s leads to inherent gaps between some packets.

Data Stream Characteristics for Continuous Media* (10)

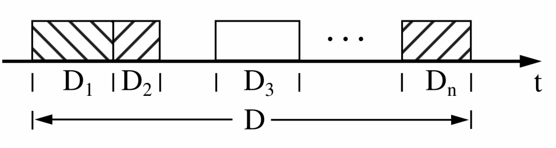


Figure 10: Non-interrelated data stream; there are gaps between packets

Example

A video signal is sampled by a camera and digitized in a computer without compression. The resulting data stream is strongly periodic, strongly regular, and interrelated, as shown in Figure 6. There are no gaps between packets. If the MPEG compression method is applied, the data stream becomes weakly periodic and weakly regular over a longer duration. When transmitted over a 16 Mbit/s token-ring network, the data stream would be non-interrelated due to the network's characteristics.

Table of Contents

Information Units (1)

Continuous media consists of a time-dependent sequence of individual information units, which are referred to as Logical Data Units (LDUs). These LDUs can be categorized as either Closed LDUs or Open LDUs based on their relationship with time:

1. **Closed LDU:** A Closed LDU has a predefined duration. This type of LDU is common in scenarios where the data is already organized into fixed durations. For example, in a computer system, audio samples are often grouped into closed LDUs, where each unit represents a specific time interval.
2. **Open LDU:** An Open LDU does not have a known duration in advance. This is typical for data streams that are generated or captured in real-time. For example, a data stream sent from a camera or microphone to a computer is considered an open LDU, as the duration of each unit is not predefined.

Information Units (2)

Granularity refers to the hierarchical decomposition of an audio or video stream into its components. In the examples above, we see that extensive information units can be described using both a symphony and motion video.

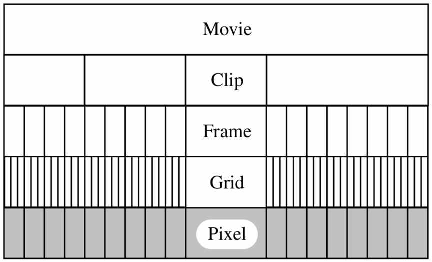


Figure 11: Granularity of a motion video sequence showing its logical data units (LDUs)

Information Units (3)

The most common examples of LDUs in continuous media are symphonies and movies, where the entire work or various components of the work can be considered as an LDU. In the case of a symphony, the LDU could be the entire composition, individual movements, specific musical scores, groups of samples (such as those grouped into 1/75-second durations), or even individual samples themselves.

Similarly, in the case of an uncompressed video sequence, the LDU could refer to the entire video clip, specific scenes, or individual frames. Each scene may consist of a sequence of images, and each image can be divided into smaller regions, such as groups of pixels. Each pixel contains luminance and chrominance values, meaning that in a video stream, the LDU could be the entire scene, a single image, or even a specific pixel within an image.

Information Units (4)

In motion video, where compression methods such as MPEG are used, inter-frame techniques are applied to reduce the data by encoding the differences between images, and this can make the LDU smaller and more efficient for transmission.

R. Steinmetz and K. Nahrstedt.
Multimedia: Computing, Communications and Applications.
Pearson, 2012.

Ralf Steinmetz and Klara Nahrstedt.
Multimedia fundamentals, Volume 1: Media coding and content processing.
Pearson Education, 2002.