Principles of Multimedia

THIRD EDITION

Ranjan Parekh

CRC Press
Taylor & Francis Group

Principles of Multimedia

Principles of Multimedia introduces and explains the theoretical concepts related to the representation, storage, compression, transmission and processing of various multimedia components, including text, image, graphics, audio, video and animation, as well as their use across various applications. The book provides the necessary programming tools and analysis technique concepts to perform practical processing tasks in software labs and to solve numerical problems at the postgraduate level. For this new third edition, every chapter has been updated and the book has been carefully streamlined throughout.

Chapter 1 provides an overview of multimedia technology, including the definition, major characteristics, hardware, software, standards, technologies and relevant theorems with mathematical formulations. Chapter 2 covers text, including digital text representations, text editing and processing tools, text application areas and text file formats. Chapter 3 explores digital image input and output systems, image editing and processing tools, image application areas, image color management and image file formats. Chapter 4 discusses 2D and 3D graphics algorithms, transformation matrices, splines, fractals, vectors, projection application areas and graphics file formats. Chapter 5 covers audio, including digital audio input and output systems, audio editing and processing tools, audio application areas and audio file formats. Chapter 6 looks at video, including digital video input and output systems, video editing and processing tools, video application areas and video file formats. Chapter 7 focuses on animation, covering 2D and 3D animation algorithms, interpolations, modeling, texture mapping, lights, illumination models, camera, rendering, application areas and animation file formats. Finally, Chapter 8 covers compression, including lossless and lossy compression techniques, and various algorithms related to text image audio and video compression. Every chapter includes solved numerical problems, coding examples and references for further reading.

Including theoretical explanations, mathematical formulations, solved numerical problems and coding examples throughout, ***Principles of Multimedia*** is an ideal textbook for graduate and postgraduate students studying courses on image processing, speech and language processing, signal processing, video object detection and tracking, graphic design and modeling and related multimedia technologies.

Principles of Multimedia

Third Edition

Ranjan Parekh

CRC Press
Taylor & Francis Group
Boca Raton New York London

CRC Press is an imprint of the
Taylor & Francis Group, an **informa** business

Designed cover image: Ranjan Parekh

Third edition published 2025
by CRC Press
2385 NW Executive Center Drive, Suite 320, Boca Raton FL 33431

and by CRC Press
4 Park Square, Milton Park, Abingdon, Oxon, OX14 4RN

CRC Press is an imprint of Taylor & Francis Group, LLC

First edition published by McGraw-Hill 2006

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
Names: Parekh, Ranjan, author.
Title: Principles of multimedia / Ranjan Parekh.
Description: 3rd edition. | Boca Raton ; London : CRC Press, 2025. |
Includes bibliographical references and index. | Contents: Text – Image – Graphics – Audio – Video – Animation – Compression. |
Identifiers: LCCN 2024029165 | ISBN 9781032643328 (hardback) | ISBN 9781032642949 (paperback) | ISBN 9781032643342 (ebook)
Subjects: LCSH: Multimedia systems. | Digital video. | Image processing–Digital techniques. | Sound–Recording and reproducing–Digital techniques. | MATLAB.
Classification: LCC QA76.575 .P345 2025 | DDC 006.7–dc23/eng/20240910
LC record available at https://lccn.loc.gov/2024029165

ISBN: 978-1-032-96109-5 (hbk)
ISBN: 978-1-032-96110-1 (pbk)
ISBN: 978-1-003-58803-0 (ebk)

DOI: 10.1201/9781003588030

Typeset in Palatino
by SPi Technologies India Pvt Ltd (Straive)

Dedicated to the Loving memory of my mother Uma Parekh (1936–2023)

Contents

Preface

I am pleased to bring out the 3rd edition of the *Principles of Multimedia* after a gap of ten years. In this edition attempts have been made to include the latest updates and standards that have come up over the past decade in the field of multimedia and at the same time reduce or remove concepts and applications that have become obsolete. Each chapter begins with a list of major key terms to give the reader a feel of a snapshot of the important ideas discussed in the chapter. These terms are expanded in the remaining sections of the chapter in their proper contexts. This edition comprises eight chapters, six for the media components and one each for the overview and compression technologies. Each chapter now contains a section on legal and ethical considerations, and another on the challenges and emerging trends. Numerical problems and MATLAB® coding are now discussed in dedicated sections at the end of the chapters. Also each chapter now contains its own reference section.

The new topics and sections added in this edition are outlined below:

Chapter 1: Multimedia – An Overview

Historical evolution of multimedia, hardware and software related to AR and VR, standalone and Web-based editing and authoring tools, APIs and libraries, standards and standards organizations, expanded lists of JPEG and MPEG standards, multimedia frontend and backend technologies, multimedia platforms, transport mechanisms, multimedia games and interactivity, multimedia content creation workflow, discrete wavelet transforms, legal and ethical considerations for multimedia products, challenges and emerging trends.

Chapter 2: Text

Historical evolution of digital text editing, details on Unicode character set, text extraction and transformation, text enhancement and effects, typography and fonts, Web-based design and applications including HTML, CSS, JavaScript, DHTML and SVG, text-to-speech (TTS) and speech-to-text (STT) applications, text localization and text internationalization, text in AR and VR applications, legal and ethical considerations in text processing, challenges and emerging trends.

Chapter 3: Image

Historical evolution of digital image processing and image processing applications like enhancement restoration segmentation, object detection in images, image analysis recognition and classification, legal and ethical considerations in image processing, challenges and emerging trends.

Chapter 4: Graphics

Historical evolution of 2D and 3D graphics, spline properties like length area and centroid, 3D rotation parallel to primary axes, 2D and 3D vectors, homogeneous vector equations, foreshortening factors, isometric diametric and trimetric projections, cavalier and cabinet projections, legal and ethical considerations in graphics production, challenges and emerging trends.

Chapter 5: Audio

Historical evolution of digital audio processing, progressive waves, key features of MIDI 2.0, audio segmentation, noise gate, dynamic range compression, legal and ethical considerations in audio production, challenges and emerging trends.

Chapter 6: Video

Historical evolution of digital video processing, video-based applications like blob detection, human detection, face detection and motion tracking, legal and ethical considerations in video production, challenges and emerging trends.

Chapter 7: Animation

Historical evolution of 2D and 3D animation, 3D surface equations, surface color and texture mapping, lights and illumination models, legal and ethical considerations in animation production, challenges and emerging trends.

Chapter 8: Compression

Historical evolution of compression techniques, LZ77, LZ78, LZW, quadtree segmentation, fractal compression, H.265 / HEVC.

In general, the book, like its previous editions, explains the concepts and principles of multimedia information processing and its applications, and, at the same time, provides hands-on techniques for solving numerical problems and writing program codes. The book has been written for graduate and postgraduate students studying courses on image, audio and video processing as well as 2D and 3D graphics modeling and animation. Objectives of the book include providing overviews of hardware and software technologies related to multimedia processing together with steps for problem-solving and practical implementations using data analysis and visualization tools.

As in previous editions, MATLAB has been used to illustrate the coding examples, primarily because of the extensive libraries of functions and toolkits supported by MATLAB which helps to decrease coding time and improve efficiency. A second reason is the benefit of the advanced visualization tools offered by MATLAB in the form of various types of 2D and 3D graphs, plots and figures for displaying image, graphics and video media components as well as playback of sample-based audio and MIDI notes. The number of MATLAB functions discussed has been increased significantly in this edition to more than 200 and a one-line description for each has been added where the functions are introduced for the first time as well as in the Appendix.

The book contains more than 90 solved numerical problems and more than 130 coding examples in MATLAB for hands-on practice in order to gain a deeper understanding of the subject matter and improve trouble shooting skills. More than 300 figures, illustrations, graphs and plots have been included to provide adequate pictorial representations and visualization aids to help understand the core concepts, especially those involving 3D environments. More than 170 references have been included for providing the sources of the information included in the book as well as for providing further reading materials to the interested reader. Wherever possible, website addresses have been mentioned to enable further studies and exploration regarding the latest updates and state-of-the-art advancements in the field. Queries and feedback can be addressed to the author via the e-mail ID author.principlesofmultimedia@yahoo.com, as well as to obtain a softcopy of the MATLAB codes on production of proof of purchase of the book.

Ranjan Parekh, PhD
Former Professor
Jadavpur University
Calcutta, India
2024

Author Biography

Dr. Ranjan Parekh, PhD (Engineering), is Former Professor, School of Education Technology, Jadavpur University, Calcutta, India. He teaches courses related to graphics and multimedia at postgraduate and doctorate levels, and his research interests include multimedia information processing, pattern recognition and computer vision.

He is also the author of the following books :

Fundamentals of Graphics using MATLAB (CRC Press, 2020) [https://www.routledge.com/Fundamentals-of-Graphics-Using-MATLAB/Parekh/p/book/9780367184827]

Fundamentals of Image Audio and Video Processing using MATLAB with Applications to Pattern Recognition (CRC Press, 2021) [https://www.routledge.com/Fundamentals-of-Image-Audio-and-Video-Processing-Using-MATLAB-With/Parekh/p/book/9780367895242]

1

Multimedia: An Overview

1.1 Multimedia Fundamentals

1.1.1 Key Terms

Hello Readers! We begin the book with a snapshot of the important terms relevant to multimedia technology, to present the contents of the entire chapter in a nutshell. These terms will be expanded in the following sections of the chapter in their proper contexts.

AR: Augmented reality (AR) is a technology that overlays digital information, such as images, text, or 3D models, onto the real-world environment, enhancing the user's perception and interaction with their surroundings. AR integrates computer-generated content seamlessly into the user's physical environment, offering an immersive and interactive experience that blurs the line between the digital and real worlds.

Artifacts: Unintended and undesirable alterations or distortions in digital images caused by compression, manipulations or other factors. These can include visual anomalies like blurriness, pixelation or color distortions, introduced during various types of signal processing.

CDN: A Content Delivery Network (CDN) is a network of distributed servers strategically positioned to deliver Web content, such as images, videos and scripts, to users based on their geographical locations. By caching and serving content from servers closer to the end-users, CDNs enhance website performance, reduce latency and improve overall user experience.

CMS: A Content Management System (CMS) is a software application that enables users to create, manage and modify digital content on a website without requiring specialized technical knowledge. CMS platforms provide a user-friendly interface for content creation, editing and organization, facilitating efficient website management for individuals and organizations.

CSS: Cascading Style Sheets (CSS) is a stylesheet language used to describe the presentation and formatting of a document written in HTML or XML. It enables Web developers to control the layout, design and appearance of Web pages by defining styles for elements such as fonts, colors, spacing and positioning.

DAM: Digital Asset Management (DAM) is a system that organizes, stores and retrieves digital assets such as images, videos and documents in a centralized and efficient manner. DAM solutions streamline the management and distribution of digital content, providing tools for version control, metadata tagging and collaborative workflows to enhance the overall accessibility and usability of digital assets within an organization.

DASH: Dynamic Adaptive Streaming over HTTP (DASH) is a standard for delivering multimedia content, such as videos, over the internet by dynamically adjusting the quality of the content based on the viewer's network conditions. DASH enables a smoother streaming experience by allowing the client device to adaptively switch between different quality levels of the video, optimizing playback for varying bandwidth and device capabilities.

DBMS: A Database Management System (DBMS) is software that facilitates the creation, organization, retrieval and management of data in a database. It provides an interface for users and applications to interact with the database, ensuring data integrity, security and efficient data handling operations.

DCT: Discrete Cosine Transform (DCT) is a mathematical operation commonly used in signal processing and image compression to transform spatial data into frequency components. It's widely employed in various multimedia applications, such as JPEG compression for images and MPEG compression for videos, helping reduce data redundancy and achieve more efficient storage or transmission.

DFT: The Discrete Fourier Transform (DFT) is a mathematical operation used to analyze

DOI: 10.1201/9781003588030-1

the frequency content of discrete signals or sequences. By converting a sequence of time-domain data points into its equivalent frequency-domain representation, the DFT is a fundamental process in fields like signal processing and telecommunications for analyzing and manipulating signals in the frequency domain.

DMCA: The Digital Millennium Copyright Act (DMCA) is a U.S. law enacted to address copyright challenges arising from digital technologies. It provides a framework for copyright holders to protect their content online and outlines procedures for service providers to respond to claims of copyright infringement.

DRM: Digital Rights Management (DRM) is a technology used to protect and control access to digital content, such as music, videos and e-books, by imposing restrictions on copying, sharing and modifying the content. DRM systems aim to prevent unauthorized distribution and ensure that users comply with licensing agreements, often through encryption and access control mechanisms.

DWT: The Discrete Wavelet Transform (DWT) is a mathematical operation to transform and analyze signals or images into different frequency components using small wave-like oscillations called wavelets as basis functions. DWT is used in signal processing and image compression and represents visual data in multiple resolutions by capturing both high and low-frequency information.

Euler Formula: Euler's formula on complex numbers is a mathematical identity that relates the exponential function, the imaginary unit (*i*), Euler's number (e), and the mathematical constant pi (π). The theorem is given by the formula: $e^{i\pi} + 1 = 0$, expressing a profound connection between fundamental mathematical constants and complex numbers on the complex plane.

Fourier Theorem: Fourier's theorem states that any periodic function can be represented as the sum of sinusoidal functions (sine and cosine) with different frequencies and amplitudes. This mathematical concept is fundamental in signal processing, allowing complex waveforms to be analyzed and synthesized using simpler sinusoidal components.

HLS: HTTP Live Streaming (HLS) is a streaming protocol developed by Apple for delivering live and on-demand multimedia content over the internet. It uses standard HTTP protocols, allowing adaptive streaming of media content, dynamically adjusting the quality based on the viewer's network conditions for a smoother viewing experience.

HTML5: HTML5 is the latest version of the Hyper Text Markup Language, widely used for structuring and presenting content on the World Wide Web. It introduces new features and improvements, such as native support for audio and video embedding, canvas elements for graphics rendering, and enhanced semantic tags, contributing to a more dynamic and interactive Web experience.

Interactivity: Interactivity in multimedia refers to the capability of engaging the user through dynamic and responsive elements, allowing them to actively participate and influence the content. This can include interactive features such as clickable buttons, animations, quizzes or user-controlled navigation, enhancing the user experience and enabling a more engaging and personalized interaction with multimedia content.

JPEG 1: JPEG (Joint Photographic Experts Group) is a widely used image compression standard that employs lossy compression techniques to reduce the file size of digital images. JPEG 1, the initial version of the standard, was introduced in 1992, defining the baseline compression algorithm using DCT-based methods and encoding method for storing photographic images with good compression efficiency.

JPEG 2000: JPEG 2000 is an image compression standard introduced by the Joint Photographic Experts Group (JPEG) in the year 2000 which employs more advanced DWT-based compression techniques compared to the original JPEG which uses DCT. It enables signals to be represented using multiples resolutions and finds extensive applications in image compression and noise reduction.

MPEG 1: MPEG (Moving Picture Experts Group) 1 is an audio/video compression standard developed to provide VHS-quality video and CD-quality audio for digital storage and transmission. Introduced in 1991, MPEG-1 is widely known for its role in creating the popular MPEG-1 Audio Layer III format, commonly known as MP3, which revolutionized digital audio compression and distribution.

MPEG 2: MPEG-2 is a video compression standard designed for higher quality digital video transmission and storage. Introduced in 1995, MPEG-2 is widely used in DVDs, digital television broadcasting, and video streaming, providing improved video quality and support for interlaced video compared to its predecessor, MPEG-1.

MPEG 4: MPEG-4 is a versatile video compression standard that supports a wide range of multimedia applications, including streaming, broadcasting and interactive media. Introduced in 1999, MPEG-4 provides efficient compression while allowing for advanced features such as object-based coding, content-based interactivity and support for various multimedia formats, making it widely utilized in online video, mobile devices, and multimedia streaming.

MPEG 7: MPEG-7 is a multimedia content description standard that focuses on providing a standardized set of descriptors to describe multimedia content for search and retrieval purposes. Introduced in 2001, MPEG-7 aims to enable efficient indexing and searching of multimedia content, offering a standardized way to describe the visual, auditory and multimedia aspects of the data.

MPEG 21: MPEG-21 is a multimedia framework standard designed to enable transparent and augmented use of multimedia resources across various networks and devices. Introduced in 2001, MPEG-21 provides a comprehensive set of tools and standards to address content creation, delivery, consumption and protection, supporting a wide range of applications in the evolving landscape of multimedia technologies.

Non-Linearity: Non-linearity in multimedia refers to the ability for users to navigate and interact with content in a customized fashion, rather than viewing it linearly from begin to end. It allows them to choose different navigational paths and experiences within the multimedia content, enabling interactive storytelling, user-controlled exploration, simulations and feedback schemes.

Nyquist Theorem: Nyquist's theorem, also known as the Nyquist-Shannon sampling theorem, states that to accurately reconstruct a continuous signal, it must be sampled at a rate that is at least twice the highest frequency present in the signal, during digitization. This theorem is fundamental in digital signal processing and ensures that sampled signals can be faithfully reconstructed without loss of information.

SNR: Signal-to-Noise Ratio (SNR) is a measure used in signal processing to quantify the ratio of the strength of a signal to the level of background noise present. A higher SNR indicates a clearer and more reliable signal, while a lower SNR implies that the signal may be more challenging to discern from the surrounding noise.

SVG: Scalable Vector Graphics (SVG) is an XML-based file format used for describing vector graphics, which can be scaled to different sizes without losing quality. SVG is widely used on the Web for creating interactive and resolution-independent graphics, offering a versatile and lightweight format for displaying diagrams, figures and illustrations.

VR: Virtual Reality (VR) is a computer-generated simulation of an immersive environment that can be interacted with and explored by users. It typically involves the use of specialized hardware, such as headsets and controllers, to provide a sensory-rich experience that can simulate real or imaginary worlds.

WebGL: WebGL (Web Graphics Library) is a JavaScript API that enables rendering 3D and 2D graphics within Web browsers without the need for plug-ins. It allows developers to leverage the GPU (Graphics Processing Unit) to create high-performance interactive graphics directly in the browser, facilitating the development of sophisticated Web-based visualizations and games.

WebRTC: WebRTC (Web Real-Time Communication) is a free, open-source project that provides real-time communication capabilities directly in Web browsers through simple application programming interfaces (APIs). It allows for peer-to-peer communication, including audio and video calling, as well as data sharing, without the need for plug-ins or additional software installations.

WebVR: WebVR (Web Virtual Reality) is an API that allows developers to create virtual reality experiences that can be accessed directly through Web browsers. It enables users to explore VR content on compatible devices, fostering accessibility and ease of use for virtual reality applications on the Web.

WebXR: WebXR (Web Extended Reality) is an API that expands on WebVR to support a broader range of immersive technologies, including augmented reality (AR) and mixed reality (MR) experiences in addition to virtual reality (VR). It enables developers to create cross-platform, immersive Web-based experiences that can be accessed on various devices, from VR headsets to AR-enabled smartphones.

1.1.2 Definitions

The word "multimedia" comes from the Latin words "multus", which means numerous, and "media", which means middle or center. More recently, the word "media" (singular "medium") started to convey the sense "intermediary." "Multimedia" therefore means "multiple intermediaries" or "multiple means." The word "multimedia" may be used as a noun (e.g., multimedia is a new technological field) and as an adjective (e.g., a multimedia document). Multimedia, in the general sense, therefore, means "multiple intermediaries" between the source and sink of information or "multiple means" by which information is stored, transmitted, presented or perceived. In modern times we use six different means to convey information: text (e.g., book chapter, newspaper article), image (e.g., photograph, scanned document), graphics (e.g., sketch, drawing), sound (e.g., speech, music), video (e.g., TV program, movie clip), animation (e.g., cartoon, simulation).

Here are some definitions from extant literature:

> Multimedia refers to the integration of various media types, such as text, images, audio, video, and interactive elements, to convey information or entertainment through digital platforms.
>
> **(Valdez, 2006)**

> Multimedia is a technology that combines different forms of media content, enabling the presentation of information in a rich, interactive format, often delivered via computer systems.
>
> **(Sharda et al., 2018)**

> Multimedia encompasses the use of multiple communication channels, including visual, auditory, and interactive elements, to engage and inform audiences, often through digital media and technology.
>
> **(Steinmetz & Nahrstedt, 1995)**

> Multimedia involves the seamless integration of text, graphics, audio, video, and interactivity to create a dynamic, immersive user experience, commonly facilitated by computers and the internet.
>
> **(Vaughan, 2014)**

> Multimedia denotes the convergence of diverse media forms, allowing the simultaneous presentation of visual, auditory, and interactive content to convey messages, educate, entertain, or enhance user experiences in digital environments.
>
> **(Moreno, 2007)**

From the above we can summarize as follows: the term "multimedia" refers to an integrated multidisciplinary field which combines various media elements including text, image, graphics, audio, video, animation and interactivity to convey information.

1.1.3 Historical Evolution

The historical evolution of multimedia spans over several decades marked by significant technological advancements and innovations. Some of the key milestones are listed below:

1960s: The foundation for multimedia was laid with the development of early computer systems and graphical user interfaces. Ivan Sutherland's Sketchpad (Sutherland, 1963) is considered one of the earliest examples of interactive computer graphics.

1970s: Witnessed the emergence of computer-based interactive multimedia with the advent of laser discs and the development of text-based computer games. The first arcade video game, "Pong", was released in 1972, introducing interactive audio-visual entertainment. Alan Kay's work on the Dynabook concept laid the groundwork for portable, multimedia-enabled computing.

1980s: Marked the rise of personal computers, which brought multimedia capabilities to homes and offices. In 1984, Apple introduced the Macintosh, which featured a graphical user interface and could handle multimedia applications. The introduction of CD-ROMs (Compact Disc-Read Only Memory) allowed for the storage and distribution of multimedia content, including encyclopedias and educational software. MIDI (Musical Instrument Digital Interface) technology facilitated digital music composition and playback.

1990s: Turning point for multimedia, with the proliferation of multimedia software, games and educational CD-ROMs. The World Wide Web was born in 1991 (Berners-Lee, 1999) bringing multimedia to the internet. HTML (Hyper Text Markup Language) allowed for the integration of text, images and hyperlinks. The release of Adobe Flash in 1996 revolutionized Web multimedia, enabling animations, games and interactive content. Digital image and video compression standards, such as JPEG and MPEG, made it possible to view photographic images and videos online.

2000s: Witnessed the convergence of multimedia technologies on mobile devices, such as smartphones and tablets. YouTube, founded in 2005, became a platform for sharing and consuming multimedia content, including videos and music. High-definition (HD) and Blu-ray Disc technology provided superior multimedia experiences, with sharper images and better sound quality.

2010s: Saw the rise of streaming platforms like Netflix and Spotify, which delivered multimedia content over the internet, leading to the decline of physical media. Virtual reality (VR) and augmented reality (AR) gained momentum, offering immersive multimedia experiences. The development of 4K and 8K video resolutions enhanced the quality of multimedia content. Interactive storytelling and narrative-driven video games reached new heights.

2020s: Continue to see the evolution of multimedia with advancements in AI-driven content creation, 5G connectivity for seamless streaming and the growth of live streaming platforms.

1.1.4 Main Characteristics

To create a successful multimedia production it is important to incorporate some of the main characteristics as discussed below.

1.1.4.1 Multiple Media

The first characteristic of a multimedia presentation is, of course, the use of multiple media. Text has been the main mode of communication for many years during the pre-multimedia era, and still continues to be one, but it is now more and more supplemented by other media, which often prove more effective. Everybody has heard the phrase: “A picture is worth a thousand words.” Going along those lines, a video must be worth a million. As computer technology progressed, in addition to text, pictures sound video and animation also started being used to communicate ideas. Pictures were subdivided into two types: the real-world pictures captured by a camera, called images, and the hand-drawn pictures like sketches, diagrams and portraits, called graphics. Text, images and graphics are together referred to as static elements, as they do not change over time. The usage of dynamic or time-varying elements like audio, video and animation began with further improvement in technology. Audio requires speakers to be played back and can comprise music, human speech or other environmental sounds like laughter, applause, horns, thunder and so on. Videos are combinations of images and sound; instead of a single image, a video is composed of a series of images shown one after another at a certain speed, while the sound is played together in the background in a synchronized manner with the visual content. Videos are referred to as motion picture, when they are recorded on film/celluloid and as motion video when recorded on electro-magnetic media. Animation is a combination of graphics and sound, a series of graphic pictures shown in a specific sequence along with the added sound. According to general convention a legitimate multimedia presentation should contain at least one static media, like text, image or graphics, and at least one dynamic media, like audio, video or animation, along with some form of interactivity and non-linearity. Different types of information might warrant the use of different types of media. For example, to show the exterior shape, size and color of a car, one would use photographic images; to show the internal parts of its combustion engine, one would prefer graphics with specific parts highlighted or detailed out if required; to depict how the pistons of its engine move or how the gearing system works, one would prefer animation; to exhibit the quality of its music system, one would use audio; and to give a sense of its pickup and braking systems, one would use video. In general, one should use static media like images and graphics if they are sufficient for explaining relevant concepts, (e.g., for depicting size, shape, color of objects) and use video only when images are not sufficient for the purpose, as video takes up a lot of space and their use should be justified (e.g., when explaining dynamic processes or moving parts). Animation is used in situations where capturing video is not practicable, e.g., when the objects are too small to be photographed (e.g., motions of sub-atomic particles), or too large (e.g., planetary bodies), or something difficult to photograph (e.g., electromagnetic waves). Text is usually used as bulleted points to highlight main concepts, while other media are used as supplementary information for better cognition.

1.1.4.2 Non-Linearity

Non-linearity indicates the capability of navigating from one point within a presentation to another point without appreciable delay. TV shows and motion pictures are considered linear presentations because the user has to watch the information being displayed, in a pre-defined sequential manner as determined by the producer or creator of the show. The sequence of frames or the timing between them cannot be changed if desired. Devices like the video cassette player allows some flexibility during playback of a recorded video cassette by enabling forward, backward, start, stop or pause operations; however, these cannot be done instantly as the tape has to rotate forward or backward to skip intermediate frames, which causes delays in displaying non-contiguous portions. In a multimedia presentation, the user is able to instantly navigate to different parts of the presentation without appreciable delays; however, such navigation is not arbitrary but only along the pathways defined by the author or developer. The author needs to ensure that the presentation behaves in a consistent manner irrespective of the sequence or order in which it is viewed.

1.1.4.3 Interactivity

To make non-linearity a possibility, a user needs to interact with a presentation. A linear presentation is usually watched passively without changing the order or sequence in which it is displayed. For a non-linear presentation, a user will have to specify the manner of non-linear navigation, e.g., directly navigate to an area of interest instead of watching something from the beginning. Such interaction is made possible through a set of interactivity elements embedded within the presentation like buttons, menu items or hyperlinks. To enable interactivity, a presentation should be able to accept user inputs and allow changing the sequence of information to be displayed. Interaction is not only limited to changing the order of display, a user can also interact to start/stop individual media elements within a presentation, e.g., click a button to start a video sequence within the presentation or pause the background music. Interactivity is a powerful tool and has potentials to achieve much more than simply switching on/off media elements: it allows one to get involved with the presentation content, like provide feedback, answer questions in a quiz, simulate an experiment and so on. This is considered very useful for applications like training and learning. It has been observed that human beings tend to remember only 20% of what they read, 40% of what they see and hear over an extended period of time, and 80% of what they do or experience. Multimedia-based education programs like CBTs (computer-based training) and CAIs (computer-aided instructions) try to exploit this property to create effective training and learning modules with better retentivity. On one hand, these use several media components to explain a subject matter more thoroughly from various aspects, and to create a more profound influence on the human mind; on the other hand, the modules are designed so as to provide ample scopes for interactivity like practice sessions, drills and tutorials to get the learner more involved. Interactivity is considered to be one of the salient features on which the new generation e-learning tools are expected to rely for greater effectiveness.

1.1.4.4 Digital Representation

Audio and video cassette players use media recorded on magnetic tapes in analog format. Magnetic tapes are called sequential access storage devices, wherein data is recorded sequentially along the length of the tape. When a specific portion of the data is required to be played back, the portion before it needs to be skipped. This creates a delay in accessing the arbitrary portions of the data which is unsuitable for requirements like teaching and learning. Multimedia productions provide instant access to different portions of the presentation using digital computers and random access devices like flash memory, hard disks and compact discs. Multimedia presentations are therefore produced and played back using digital representations. Each of the media types, like text, image, audio and video needs to be expressed in the digital form before it can be utilized within a multimedia presentation. Digital representation has other advantages too, like software-based programs that can be used to edit the digitized media in various ways to improve their quality and compress file sizes to increase performance efficiency (Steinmetz & Nahrstedt, 2004). Digitization of the media components requires specialized hardware, software, compression schemes and file formats.

1.1.4.5 Integrity

One would tend to believe that any set of media files played on a computer would comprise a multimedia presentation. For example, we can edit a text document in a word processor, play a video clip in the Media Player and listen to an Audio-CD using the CD-Player. All these applications collectively, however, would not constitute a multimedia presentation. To clarify that, we introduce another important characteristic of a multimedia presentation called integrity. It implies that although there may be several media types present and playing simultaneously, they need to be integrated or be part of a single entity, which is the presentation. We should not be able to separate out the various media and control them independently; rather they should be controlled from within the framework of the presentation. Moreover, the presentation should decide how the individual elements can be controlled. For example, it might or might not be possible to turn off the background audio depending on whether the developer has allowed it from within the presentation.

1.1.4.6 Narrative and Storytelling

A compelling narrative or storyline is often at the heart of a multimedia presentation. Stories create a logical flow and help the audience connect with the content on an emotional level. From interactive e-books and multimedia-rich documentaries to video games and virtual reality experiences, multimedia provides a versatile canvas for crafting compelling stories. It enables creators to engage their audiences on multiple sensory levels, enhancing emotional resonance and understanding. Multimedia also offers opportunities for user interaction and personalization, allowing consumers to actively participate in the unfolding narrative. Whether in entertainment, education, or marketing, the capacity to blend various media types results in storytelling that is not only informative and entertaining but also emotionally impactful, leaving a lasting impression on audiences (Nunes et al., 2017).

1.1.4.7 Informational Value

Multimedia presentations need to be informative and convey a message or share knowledge. Content should be accurate, well researched and relevant to the audience's needs or interests. Overloading the presentation with unnecessary media should be avoided. Through a seamless fusion of text, images, audio, video and interactive elements, multimedia presents information in a comprehensive and accessible manner. This convergence of media types enables content creators to provide context, depth and clarity to their messages, catering to diverse learning and communication styles. Since information needs of audiences differ, frequently a multimedia presentation would list out the pre-requisites of the target audience for which it has been designed and the objectives it is expected to fulfill in terms of learning out comes.

1.1.4.8 Adaptability and Compatibility

Multimedia presentations should be designed to work across various devices and platforms, including computers, tablets, and smartphones. Compatibility ensures that the audience can access the content easily. Multimedia presentations should be scalable to accommodate different screen sizes and resolutions. This ensures that the content remains visually appealing and functional on various devices. Additionally, multimedia files are often encoded in standard and open formats that are compatible with various software and hardware configurations, promoting interoperability. This ensures that presentation can be easily shared, integrated into different applications and enjoyed by a broad audience without the constraints of proprietary systems.

1.1.4.9 Engagement and User Experience

An effective multimedia presentation is designed with the user in mind. It aims to engage the audience and create a positive user experience through intuitive navigation and well-structured content. Multimedia content leverages the power of visuals, audio, and interactivity to immerse users in compelling experiences. From interactive storytelling to visually rich presentations and video games, multimedia heightens engagement by appealing to multiple senses and emotions. By catering to diverse learning styles and preferences, multimedia not only educates and entertains but also enhances retention and understanding. Its capacity to evoke emotions, deliver information effectively and offer interactive elements ensures that users remain deeply engaged and invested in the content, making it an invaluable tool for communication, entertainment, and education.

1.1.4.10 Feedback and Assessment

Interactive elements can include quizzes, surveys, or feedback forms to gather information from the audience or assess their understanding. Feedback mechanisms enhance the learning process. Use of analytics tools to track audience engagement and gather insights into how users interact with the presentation can lead to improvements for future presentations. In other interactive applications, multimedia allows users to make informed decisions, play games and explore virtual environments while receiving instant feedback on their actions. By facilitating this feedback loop, it opens up scopes for future improvements that cater more closely to user needs.

1.1.4.11 Accessibility and Consistency

Accessibility ensures that content can be enjoyed and understood by a diverse audience, including individuals with disabilities. Multimedia technologies offer various tools and features, such as closed captions, screen readers and alternative text for images, to make content accessible to individuals with visual or hearing impairments. In multimedia design, consistency involves maintaining uniformity in the presentation of elements, including layout, color schemes, fonts, navigation, and user interface components. Consistency ensures that users can predict how to interact with multimedia content, making it intuitive and easy to navigate. Whether in websites, mobile apps or educational materials, a consistent design fosters a sense of familiarity, reducing cognitive load and enhancing usability. It also reinforces the brand identity and message, making content more recognizable and memorable.

1.1.4.12 Legal and Ethical Considerations

In a digital landscape where multimedia content is prevalent, issues surrounding intellectual property, copyright and fair use come to the forefront. Content creators must be mindful of using licensed materials and respecting the rights of others. Furthermore, the rise of deep-fakes and misinformation underscores the importance of ethical practices in multimedia production and distribution. Users, on the other hand, need to be discerning consumers of multimedia, understanding the potential for misinformation and respecting privacy rights.

1.2 Application Areas

Multimedia has found extensive application in various and varied fields. Text, image, graphics, audio, video and animation can be integrated together to provide a wealth of resources and multisensory user experiences in home entertainment and gaming, education and learning, business and industry, advertising and marketing, communication and marketing, engineering and manufacturing, medicine and healthcare along with a host of other application areas. In recent years, AR (augmented reality) and VR (virtual reality) have revolutionized how we perceive and interact with the digital world. AR enhances our real-world environment by overlaying digital information and objects, making it invaluable in areas like gaming, navigation and industrial training. On the other hand, VR immerses users in entirely virtual environments, offering unparalleled experiences in fields such as gaming, education and therapy (Chen & Fragomeni, 2020). Some of the major applications include, but not limited to, the following.

1.2.1 Home Entertainment

Application of multimedia technology related to home entertainment includes computer-based games, interactive encyclopedias, storytelling, cartoons, movies and other immersive experiences. Through high-definition displays, surround sound systems and streaming services people can enjoy all the above experiences from the comfort of their homes. Audio/video-on-demand systems employing multimedia technologies enable users to specify the content they would like to watch, not simply by their names, but by a host of other criteria like creator, music, genre, year of release, and awards. Smart TVs integrate multimedia content with internet connectivity offering access to a vast array of streaming platforms and online content. Additionally virtual reality headsets bring a new dimension to home entertainment and games, allowing users to experience immersive games and immersive storytelling (Steinkuehler et al., 2012).

1.2.2 Education and Learning

Multimedia technology enhances education by providing dynamic and engaging learning experiences. Education-based multimedia applications include learning packages, CBTs (computer-based training), CAIs (computer-aided instructions) and simulation of lab experiments (especially those which cannot be easily performed in practice). Different aspects of the course curriculum which cannot be explained or grasped easily through simple text and images (as offered by the books) could be presented through video clips, animation, 3D modeling and audio annotations, AR and VR for making them easily understandable. Interactive whiteboards, educational videos and online courses offer students access to a wealth of resources that cater to diverse learning styles. Moreover, expertise of the best instructors could be used to create these online content and distribute them to students all over the globe as part of "distance education" or "e-learning" programs. The multisensory perceptions of such study material is expected to provide a good grasp of the subject matter and interactivity elements to provide for better retention (Clark & Mayer, 2016).

1.2.3 Business and Industry

Multimedia technology has revolutionized various aspects of operations and communication in the business and industry sector. Through multimedia presentations and virtual demonstrations companies can effectively showcase products and services to clients and stakeholders, which facilitate business transactions and marketing endeavors. Computer-based learning and training packages for employees, both technical and marketing, enable organizations to maintain a high level of staff training and development. Many people can use each of these courses when imparted using networked environments and do not need to spend time away from office. Also people can learn at their own paces without the need to hire full time instructors. Because the best instructors could be used to make these materials, they could be of a high quality, to provide hands-on learning experiences enhancing skill development and safety protocols. Additionally multimedia technology enables remote collaboration through video conferencing and virtual meetings facilitating seamless communication among team members regardless of geographical location (Jong & Sarti, 2012). Business houses utilize multimedia technology to display items difficult to stock like glassware, chemicals and industrial equipment to prospective customers. Real-estate agents can display interior and exterior of buildings along with necessary information like dimension and price. In electronic commerce, customers can compare different products in relation to their quality, price and appearance without leaving their homes and offices. Once they decide on purchasing an item, hyperlinks could connect them directly to an online store where they could place their orders. The benefits include saving on space, inventory and distribution (Milutinovic, 2013).

1.2.4 Advertising and Marketing

Multimedia is a driving force in advertising and marketing. Interactive multimedia advertisements, video commercials and social media campaigns use graphics, video and sound to capture the audience's attention. Advertisers employ multimedia techniques to create immersive brand experiences. Multimedia advertising encompasses a wide range of formats, including video commercials, interactive banners, social media campaigns and multimedia-rich websites. Video content, in particular, has become a cornerstone of advertising strategies, providing a compelling medium to convey brand messages and showcase products or services, especially in online platforms like YouTube, which in turn provide incentives to create quality content. Interactive multimedia elements, such as clickable ads and immersive experiences, enable brands to create memorable and engaging campaigns. Online catalogs can also server as expert systems, which could take a list of criteria from customers and provide a list of products which satisfy them. This would serve to guide the customers into choosing the appropriate items they want. Such systems can also be useful in providing after-sales services by taking in a list of symptoms from a user and suggesting to them the possible actions to be taken including troubleshooting or subscribing to on-call services. Multimedia also facilitates cross-platform advertising, allowing seamless integration across television, online platforms and social media. With the widespread adoption of smartphones and tablets, multimedia technology is extensively used in mobile advertising campaign (Hua, 2010).

1.2.5 Communications and Information

Multimedia technologies facilitate creation of engaging and informative digital content for websites, blogs, social media, e-books and other digital publications. These applications allow real-time interactions between people who need to work together but cannot be in the same place at the same time. In order to cater to these needs, many public and private networks support services like voice-mail and teleconferencing. Voice-mail enhances the effectiveness of normal e-mail by allowing messages to be composed and sent using audio instead of text, while teleconferencing is used where the parties need to communicate in real time. Internet Telephony involves establishing long distance telephone calls by logging into a local server and using the Internet to connect to a remote server. Other similar uses include groupware/whiteboard applications in intranets, e-learning applications in virtual classrooms. In an information kiosk, information is accessed through a touch screen and viewed on a monitor. A wide range of information, products and services, like for example, sales product catalogs, transport timetables, way finding assistance, news updates, event schedules, ticket purchases, bill payments and so on. Links to payment systems and distribution channels would enable the kiosks to accept orders and promote e-commerce practices. Kiosks can also be used to capture statistical data for an in-depth market research to be carried out on consumer trends. Multimedia elements can be employed to convey information dynamically based on user requirements, often with audio cues, step by step instructions, customizable fonts and subtitles to cater to a wide range of users (Tatipamula & Khashnabish, 1998).

1.2.6 Engineering and Manufacturing

In computer-aided design (CAD) and manufacturing (CAM) multimedia elements such as 3D modeling, animations and simulations play a crucial role in visualizing complex structures and systems. Engineers use multimedia presentations to communicate design concepts, project proposals and technical documentation effectively. Virtual reality (VR) and augmented reality (AR) technologies, which heavily rely on multimedia, are increasingly employed in engineering for immersive design reviews and training simulations.

Multimedia presentations aid in the analysis of data from simulations, experiments and sensor networks, providing engineers with a comprehensive understanding of performance metrics. Product lifecycle management (PLM) systems utilize multimedia technology to manage and coordinate entire lifecycle of engineering products from conception and design to manufacturing maintenance and disposal. Multimedia technology enhances engineering education and training through virtual laboratories and simulation-based training modules, and facilitates visualization and analysis of sensor data and simulation results. Collaborative tools like video-conferencing and interactive whiteboards promote teamwork and project completion (Chen & Shyu, 2013).

1.2.7 Medicine and Healthcare

Multimedia technology has revolutionized the way medical information is communicated, medical procedures are conducted and patient care is administered. High-quality images, video, animation, virtual reality has led to advancements of medical education, patient communication, diagnostic tools and therapeutic interventions. Medical students and professionals can access a wealth of visual resources, including 3D anatomical models, surgical procedure simulations and interactive case studies. Virtual dissections, medical animations and multimedia presentations enhance understanding, making complex medical concepts more accessible and memorable. Modalities such as magnetic resonance imaging (MRI), computed tomography (CT) and ultrasound generate detailed visual representations of internal structures. The integration of multimedia in telemedicine facilitates remote consultations and enhances the quality of virtual healthcare interactions. Video conferencing, multimedia-rich medical records and interactive patient portals allow healthcare professionals to remotely assess patients, discuss treatment plans, and provide visual explanations of medical conditions, fostering a more personalized and accessible healthcare experience. Archives of X-ray images, CT scans and ultrasonography scans enable doctors to provide better consultations, and serve as an expert system. Transmission of these images over geographical boundaries can help in the growth of telemedicine and medical repositories. The DICOM (Digital Imaging and Communications in Medicine) standard is a foundational framework in the field of medical imaging and healthcare informatics, which not only standardizes the structure and format of medical images but also the communication and networking protocols required for seamless data exchange within healthcare facilities (Tyagi et al., 2021).

1.2.8 Travel and Tourism

Multimedia technology enables creation of virtual tours and destination marketing campaigns that showcase tourist attractions, hotels, resorts and landmarks. Virtual reality, 360° videos and interactive content provide potential customers with immersive experiences allowing them to virtually explore destinations before making travel decisions. Hotel owners can utilize the technology to display details of facilities offered at various hotels at different locations. Museums and art galleries may provide digital versions of their specimens for public viewing, exhibition, auction, etc. A multimedia system implementing intelligent travel agent software would enable the users to specify their travel needs and budgets, and then make the system list out possible destinations that meet these criteria and create multimedia presentations about one or more of these. Travels apps incorporate multimedia elements such as maps, GPS, travel reviews, destination guides to offer personalized travel itineraries. Airlines and hospitality provides can provide passengers with personalized entertainment options and enhance on-board comfort. Museums, heritage sites and cultural centers can use audio guide, digital storytelling and interactive exhibits to engage visitors about local cultures traditions and history (Pirolli, 2018).

1.2.9 News and Journalism

Multimedia technologies have allowed news organizations to present stories in more engaging and informative ways. News websites incorporate images, videos and interactive graphics to provide comprehensive coverage and context. Live streams and social media posts enable real-time reporting of events as they unfold. Multimedia also offers new avenues for citizen journalism, where individuals can capture and share news through their smartphones and social media platforms. The integration of multimedia in journalism has made news more accessible and engaging for audiences, enhancing understanding and awareness of global events. Live streaming, live blogs and social media platforms enable journalists to report news in real time providing audiences with immediate access to headlines, live updates and eyewitness accounts. Data journalism enables journalists to analyze, interpret and visualize complex datasets to uncover insights, trends and patterns in data. Mobile journalism is a growing trend, which empower journalists to capture, edit and share news content on the go (Hernandez & Rue, 2015).

1.2.10 Design and Architecture

Architects, interior designers and urban planners employ multimedia to create virtual 3D models, walk-throughs, and flyovers, offering clients a realistic and immersive understanding of proposed projects. Multimedia also facilitates collaborative design processes, allowing teams to share ideas and visualize concepts in real time, regardless of geographical distances. Furthermore, multimedia enhances architectural presentations and marketing materials, enabling professionals to showcase their work through visually stunning renderings, animations and interactive portfolios. Augmented and virtual reality technologies enable architects to immerse stakeholders in virtual environments, enhancing the understanding of the proposed designs.

1.2.11 Museums and Exhibits

Multimedia technologies have enabled museums to create dynamic and interactive exhibits that captivate audiences of all ages. Interactive touchscreens, immersive virtual reality experiences and multimedia guides have become integral to modern museums, providing deeper insights into art, history and science. Visitors can explore artifacts and artworks in greater detail, access contextual information and even participate in hands-on simulations. Additionally, multimedia has expanded the reach of museums, allowing them to create virtual exhibits and online collections accessible to a global audience (Vatsky, 2023).

1.2.12 Government and Public Services

Government agencies use multimedia to communicate policies, health information, emergency alerts and civic engagement initiatives effectively. Multimedia elements, such as video conferences and webinars, facilitate remote collaboration and public participation in government meetings and decision-making processes. The application of multimedia in government also extends to e-learning platforms for employee training and public education (Mehta & Gaie, 2023).

1.2.13 Farming and Agriculture

Multimedia technologies offer farmers valuable resources, from online tutorials and interactive applications to informative videos and webinars. These resources cover a wide range of topics, including crop management, pest control, livestock care, and sustainable farming practices. Multimedia also plays a role in precision agriculture, utilizing sensors and data visualization to improve crop yield and resource management. Moreover, multimedia assists in sharing vital agricultural knowledge and research, aiding farmers in adapting to changing environmental conditions and emerging challenges (Bansal & Uddin, 2022).

1.2.14 Social Media and Content Creation

Social media platforms like Facebook, Instagram, Twitter(X) and YouTube are powered by a diverse mix of multimedia elements, from images, videos and GIFs to audio clips and interactive posts. Content creators, influencers and individuals use multimedia to craft compelling narratives, share their experiences and connect with global audiences. Multimedia content spans a wide spectrum, including vlogs, podcasts, visually engaging info-graphics and live streaming, all fostering vibrant online communities (Adornato, 2021).

1.2.15 Sports and Athletics

High-definition video broadcasts, slow-motion replays and interactive graphics provide viewers with immersive and engaging coverage. In addition, multimedia analysis tools have become essential for coaches and athletes, enabling in-depth performance assessments, injury prevention and tactical strategies. Multimedia-driven sports journalism has also become integral, offering fans in-depth insights through podcasts, videos and interactive articles. In these platforms, multimedia fuels user-generated content, shaping trends, sparking discussions and amplifying digital voices

1.2.16 Environment and Conservation

Multimedia technologies enable researchers and environmentalists to capture and share stunning visuals of ecosystems, wildlife and natural phenomena, conveying the urgency of conservation efforts. Educational institutions and environmental organizations use multimedia resources, such as documentaries, interactive websites and virtual tours, to inform the public about critical environmental issues. These tools encourage advocacy, support sustainable practices and foster appreciation for the natural world. Additionally, GIS (Geographic Information System) mapping, which relies on multimedia data, aids in ecological research, land management and conservation planning (Gadsden & Artz, 2023).

1.2.17 Pattern Recognition and Deep Learning

In recent years multimedia technologies have played crucial roles in enhancing algorithms of pattern recognition, computer vision and Content Based Storage and Retrieval (CBSR) systems. Traditionally, data-searching activities have been performed on textual databases by string matching. As large repositories of media elements like images, audio and video are growing up all over the world, efficient methods of searching non-textual media are being developed. Machine learning algorithms trained on multimedia datasets are utilized to detect and recognize faces, fingerprints, signature, speech, speaker and other patterns in data. Deep learning models trained on large-scale datasets utilize convolutional neural networks (CNN) to classify multimedia data into predefined categories. Interpretation of hand gestures, body movement and facial expressions, have led to the development of human computer interface (HCI) systems. Biometric systems identify and verify individuals based on their facial features, fingerprints, iris patterns, voice patterns and other biometric characteristics. Video analysis and recognition systems interpret motion patterns and events in video streams to enable automated classification of sports videos, crowd management from drone videos, and scene understanding of surveillance videos. Voice activated systems recognize speech and speaker patterns from audio streams in security and authentication systems. Medical diagnosis systems use CBSR techniques for detecting diagnosing and monitoring medical conditions and diseases. Semantic segmentation algorithms use deep learning models to partition images and videos into meaningful regions like foreground objects and background areas to enable robotic systems and automated vehicular systems to function properly (Marquis & Furht, 2012).

1.3 Multimedia Hardware

To make multimedia work requires an array of hardware devices and instruments of varying functionalities. These range from input devices like cameras, scanners, microphones to processing units like CPU, GPU, DSP to storage devices like HDD, SSD, NAS to output devices like monitors, tablets, printers to network devices like hubs, routers, switches to specialized hardware like capture cards, head mounted displays and haptic gloves.

1.3.1 Input Devices

Input devices are essential for capturing and digitizing multimedia content for enabling users to interact with digital platforms. They include a wide range of devices tailored for specific tasks ranging from scanners for documents, to cameras for images and video, and microphones for audio.

1.3.1.1 Cameras

Digital Cameras: Digital cameras come in various types, including compact point-and-shoot cameras, DSLRs (digital single-lens reflex) and mirror-less cameras. They offer a wide range of features and capabilities, making them suitable for different applications. DSLRs and mirror-less cameras are favored by professional photographers and videographers for their advanced settings, interchangeable lenses and high-quality image sensors.

Camcorders: Camcorders are dedicated video cameras designed for recording high-quality video. They come with features like image stabilization and versatile zoom lenses. Professional camcorders offer advanced settings for cinematographers and videographers.

Action Cameras: Action cameras are compact, rugged cameras designed for capturing extreme sports and outdoor activities. They are known for their durability and wide-angle lenses, which provide a unique perspective for action shots and video recording.

1.3.1.2 Scanners

Flatbed Scanners: Flatbed scanners are the most common type and are suitable for scanning flat and single-sheet documents, photographs and images. They offer high-resolution scanning and color accuracy, making them ideal for archiving and digitizing printed media.

Sheet-Fed Scanners: Sheet-fed scanners are designed for scanning multiple sheets of documents quickly. They are commonly used in office environments for digitizing large volumes of documents, such as contracts, invoices and forms.

Film Scanners: Film scanners are specialized devices for digitizing photographic film, including negatives and slides. They provide high-resolution scans and color accuracy, making them a valuable tool for photographers and archivists.

1.3.1.3 Microphones

Condenser Microphones: Condenser microphones are highly sensitive and deliver exceptional audio quality. They are commonly used in studios for recording vocals, musical instruments and podcasts. Condenser microphones require external power, typically provided through a USB connection or an audio interface.

Dynamic Microphones: Dynamic microphones are robust and less sensitive to sound variations. They are ideal for live performances, broadcasting and recording instruments like drums and electric guitars. Dynamic microphones do not require external power and are known for their durability.

Lavalier Microphones: Lavalier microphones, often referred to as lapel or clip-on microphones, are small and unobtrusive. They are commonly used in broadcasting, presentations and interviews, as they can be discreetly attached to clothing for clear audio capture.

1.3.1.4 MIDI Controllers

MIDI (Musical Instrument Digital Interface) controllers are used by musicians and music producers to create, record and perform music electronically, including controlling virtual instruments and synthesizers in digital audio workstations (DAWs). MIDI controllers come in various forms, including MIDI keyboards with velocity-sensitive keys and MIDI pad controllers with pressure-sensitive pads for drum programming and performance. Many MIDI controllers feature knobs and faders that can be assigned to control parameters in music software, offering tactile and real-time manipulation of sound.

1.3.2 Digitizing Devices

Specialized hardware plays a crucial role in converting analog data to digital format to be used as input to multimedia processing systems.

1.3.2.1 Graphics Tablets

Graphics tablets, also known as pen tablets or drawing tablets, are used by graphic designers, digital artists, and illustrators for precise and natural input when creating digital artwork, graphics, and animations. Graphics tablets come with a stylus pen that offers pressure sensitivity, allowing artists to control line thickness and opacity by adjusting the pressure they apply, akin to traditional drawing and painting. Many graphics tablets feature shortcut buttons on the stylus or the tablet surface, providing quick access to frequently used functions and tools in graphic design software.

1.3.2.2 Audio Capture Cards

Audio capture cards or sound cards are essential for recording, digitizing and playback of audio in digital computers. These offer input ports for microphones and audio players, for capturing clean and clear audio signals. They provide low-latency monitoring via bundled software, allowing musicians and audio professionals to hear themselves in real time while recording, without audio delays. They also have multiple output ports for routing audio to various speakers, headphones and audio storage devices.

1.3.2.3 Video Capture Cards

Video capture cards are used to capture video from external sources like cameras, gaming consoles and video players for use in video production, live streaming and content creation. These cards offer inputs compatible with HDMI and SDI, ensuring compatibility with a wide range of video sources. Video capture cards can capture video in real time, enabling live streaming and recording without delay. Many capture cards come with built-in streaming support, allowing users to broadcast live video content directly to online platforms.

1.3.3 Processing Units

Multimedia content often requires significant processing power for tasks such as editing, rendering and real-time effects. Processing units are responsible for executing these operations, and they include a variety of components tailored for specific functions:

1.3.3.1 Central Processing Unit (CPU)

The CPU is the central processing unit of a computer and plays a crucial role in multimedia content creation and manipulation. Many CPUs feature multiple cores, allowing them to execute multiple tasks simultaneously. This is particularly valuable for multimedia applications, which often involve running complex operations in parallel. The clock speed, measured in gigahertz (GHz), determines how quickly a CPU can execute instructions. High clock speeds are beneficial for tasks that require rapid processing, such as real-time rendering and video encoding. Hyper-threading is a technology that enables each CPU core to handle multiple threads simultaneously. This feature enhances multitasking and can improve the performance of multimedia applications. Enthusiasts and professionals may overclock their CPUs to increase their clock speeds beyond the manufacturer's specifications. Overclocking can provide additional processing power for resource-intensive multimedia tasks.

1.3.3.2 Graphics Processing Unit (GPU)

GPUs are specialized processors primarily responsible for rendering graphics, but they also contribute significantly to multimedia production. GPUs excel in parallel processing, making them suitable for tasks like video rendering, 3D modeling and real-time effects. They can handle multiple operations simultaneously, resulting in faster multimedia processing. Many modern GPUs support technologies like CUDA (Compute Unified Device Architecture) for NVIDIA GPUs [www.developer.nvidia.com/cuda-zone] and OpenCL (Open Computing Language) [www.khronos.org/opencl], enabling them to accelerate multimedia and scientific applications through parallel computing. Video RAM (VRAM) is dedicated memory on a GPU used to store textures, frame buffers and other graphics-related data. A larger VRAM capacity is advantageous for rendering high-resolution textures and complex 3D scenes.

1.3.3.3 Digital Signal Processor (DSP)

DSPs are specialized processors designed for efficient audio signal processing and real-time audio effects. DSPs are commonly used in audio equipment, such as digital audio workstations (DAWs) and effects processors, to apply effects like reverb, equalization and compression to audio signals in real time. DSPs offer low-latency audio processing, making them ideal for live audio applications, including recording, mixing, and live sound reinforcement. They are optimized for specific audio tasks and are often found in audio interfaces, hardware synthesizers and effects units. They offload audio processing from the CPU, ensuring smooth real-time performance.

1.3.4 Storage Devices

Efficient storage devices are essential for preserving and managing multimedia content. A variety of storage options are available, each with specific characteristics and use cases.

1.3.4.1 Hard Disk Drives (HDD)

HDDs offer a range of storage capacities, from gigabytes to multiple terabytes, making them suitable for archiving and storing vast amounts of multimedia data, including photos, videos, and audio recordings. They are robust and can withstand regular use, making them reliable for long-term storage and less susceptible to physical damage. HDDs are cost-effective in terms of storage capacity, and provide a larger storage-to-cost ratio, making them a practical choice for multimedia professionals with substantial storage needs. HDDs contain spinning magnetic platters and read/write heads and are widely used for media libraries, backups and content archiving due to their affordability and durability. They are also used as external drives for archives.

1.3.4.2 Solid-State Drives (SSD)

SSDs are significantly faster than HDDs in terms of data access and transfer speeds. This speed advantage is critical for multimedia applications that require rapid data retrieval, such as video editing and gaming. Also SSDs do not have any moving parts, which make them less prone to physical damage and failure. Usually SSDs are robust and can withstand shocks, vibrations and temperature variations, making them suitable for portable multimedia production environments. SSDs can come in various form factors, allowing them to fit in a range of devices such as laptops, desktops and external enclosures.

1.3.4.3 Network Attached Storage (NAS)

NAS devices provide a centralized storage solution that can be accessed and shared by multiple users over a network. They are particularly valuable for collaborative multimedia projects and remote access to multimedia files. Many NAS systems support redundancy through RAID (Redundant Array of Independent Disks) configurations, which ensures data integrity and availability, even in the event of drive failure. NAS devices come with features for remote access, allowing multimedia professionals to access their content from anywhere with an internet connection. NAS systems can function as media servers, enabling the streaming of multimedia content to various devices, including smart TVs, gaming consoles and mobile devices. NAS devices often provide backup and synchronization solutions, which are crucial for preserving multimedia content and ensuring data security.

1.3.4.4 Optical Drives

Optical drives, such as CD/DVD/Blu-ray drives, are used for reading and writing optical media. They are still relevant for distributing multimedia content on physical media and archival purposes. Different types of optical discs, including CD-ROM, DVD-R and Blu-ray discs, offer varying storage capacities; Blu-ray discs, for example, can store high-definition video content. Optical media is often used for long-term archival of multimedia content, providing a physical backup of critical data. They are popular for playback and distribution of multimedia content such as movies, music albums and software installations. One advantage of optical drives over magnetic storage is that they are unaffected by strong magnetic and electric fields in the vicinity, hence they are preferred when such situations are encountered.

1.3.5 Output Devices

Output devices are essential for presenting and delivering multimedia content to the audience in various forms. They encompass a range of devices tailored for specific multimedia applications:

1.3.5.1 Visual Display Units (VDU)

Monitors: Monitors come in various sizes and resolutions to cater to different multimedia needs. High-resolution displays are crucial for tasks like video editing, graphic design and professional content creation. Professional-grade monitors often feature accurate color reproduction, making them suitable for work like photo and video editing. Some monitors offer high refresh rates, which are particularly important for gaming and video playback. High refresh rates ensure smoother motion and reduced motion blur. There are different panel types, such as CRT (cathode ray tube). LCD (liquid crystal display), LED (light emitting diode) and PDP (plasma display panel), each with unique characteristics that impact factors like color accuracy, response time and viewing angles.

Touchscreen Displays: Touchscreen displays are used in interactive multimedia applications and kiosks, allowing users to engage directly with multimedia content like product information, transport timetables, tourist guides and stock exchange fluctuations.

Projectors: Projectors are used to display multimedia content on a larger scale. They are valuable for presentations, home theaters and events. High-end projectors offer features like 4K resolution and 3D support.

1.3.5.2 Printers

Inkjet Printers: Inkjet printers are versatile and suitable for printing high-quality color images and graphics. They are often used for photo printing and graphic design projects.

LASER Printers: Laser printers offer fast and efficient printing of text documents and are commonly used in offices. Color laser printers are used for producing marketing materials and graphics.

Large-Format Printers: Large-format printers are designed for printing posters, banners and high-resolution graphics in sizes larger than standard letter or A4 paper.

3D Printers: 3D printers are used for creating three-dimensional objects from digital designs. They have applications in product prototyping, modeling and even custom multimedia accessories.

1.3.5.3 Speakers and Headphones

High-quality speakers are essential for playing audio content with fidelity. They come in various configurations, including 2.0 (stereo), 2.1 (stereo with a subwoofer), and 5.1/7.1 (surround sound) setups. Headphones are critical for tasks like audio editing, sound design and music production. Open-back headphones are favored for their natural sound reproduction, while closed-back headphones provide sound isolation in noisy environments. Soundbars are compact speaker systems designed for home theater setups and provide improved audio quality for multimedia content consumption. Wireless audio systems enable the streaming of audio content to speakers or headphones via Bluetooth or Wi-Fi, offering convenience and mobility. Virtual surround sound technologies create a spatial audio experience through headphones, simulating a multi-speaker setup.

1.3.6 Network Infrastructure

Network infrastructure plays an important role in multimedia content distribution, streaming, collaboration and real-time communication. It encompasses a variety of components and technologies.

1.3.6.1 Routers and Switches

Home routers are the central networking devices in residential environments, which connect multiple devices to the internet and each other, allowing multimedia content sharing and streaming within a home network. In business and large-scale multimedia production environments, enterprise-grade routers are employed to manage network traffic, ensure security and provide high-speed connectivity.

Ethernet switches are used to connect multiple devices within a local area network (LAN) and facilitate high-speed data transfer between multimedia workstations, servers and storage devices, enhancing collaboration and data sharing. In professional multimedia production settings, managed switches offer greater control over network traffic, including the ability to prioritize multimedia data to ensure uninterrupted streaming and real-time collaboration.

1.3.6.2 Wireless Access Points (WAPs)

Wireless access points provide Wi-Fi connectivity, allowing multimedia professionals to work wirelessly with their devices. This is particularly valuable in collaborative environments, studios and locations where wired connections are impractical. Wi-Fi standards, like 802.11ax/Wi-Fi 6, offer high-speed wireless connections suitable for multimedia applications and enhanced performance, especially in crowded areas. WAPs work by broadcasting wireless signals within a specific range allowing devices to connect to the network seamlessly. Additionally WAPs can support technologies like MIMO (multiple input multiple output) and beamforming (focusing transmitted signal directly to intended receiver) to enhance signal strength and coverage.

1.3.6.3 CDN and QoS

Content Delivery Network (CDN) is a critical part of network infrastructure for online multimedia content delivery and distribution of multimedia content to servers strategically located around the world, reducing latency and ensuring fast delivery to users regardless of their geographical location. CDNs employ load balancing to distribute traffic efficiently, ensuring multimedia content can handle sudden surges in demand, cache content in servers, optimize bandwidth usage and reduce network congestion. Moreover, CDNs often include security features to protect multimedia content from unauthorized access. CDNs improve Quality of Service (QoS) by reducing latency, increasing throughput and enhancing reliability for end users. QoS is a network technology that prioritizes certain types of data traffic to ensure a consistent and high-quality experience for multimedia applications. QoS ensures that video conferences and real-time collaboration tools receive priority, resulting in reduced lag and smooth communication. Streaming media, such as high-definition video and audio, benefits from QoS to prevent buffering and maintain playback quality. Voice over Internet Protocol (VoIP) calls and multimedia content with voice components are less prone to interruptions and delays with QoS.

1.3.7 AR and VR Hardware

1.3.7.1 AR (Augmented Reality) Hardware

Smart Glasses: AR smart glasses, like Microsoft HoloLens and Google Glass, overlay digital information onto the user's view of the real world. Key features include:

Smartphone AR: Many smartphones are equipped with AR capabilities, enabling users to experience AR through their phone's camera and screen. ARKit for iOS [developer.apple.com/augmented-reality/arkit] and ARCore for Android [developers.google.com/ar/develop] are software frameworks that support AR on mobile devices.

Wearable AR: Some AR wearable, like the Magic Leap [www.magicleap.com], are designed for more immersive AR experiences by integrating sensors, cameras, and spatial audio.

Industrial and Medical AR: AR hardware is used in industrial settings for remote assistance and training. In the medical field, AR glasses can display patient data and surgical guidance to healthcare professionals.

Immersive AR: Immersive AR experiences are delivered through headsets that offer a wide field of view and realistic 3D interactions, blurring the line between AR and VR.

1.3.7.2 VR (Virtual Reality) Hardware

Head-Mounted Displays (HMDs): VR HMDs are the primary hardware for VR experiences. They are designed to be worn on the head and cover the user's field of vision with stereoscopic displays.

Controllers: VR controllers are handheld devices that allow users to interact with the virtual environment. They may include buttons, triggers and motion sensors for precise tracking and manipulation.

Room-Scale Tracking: Room-scale tracking systems use external sensors or cameras to track the user's movement in physical space, allowing for more natural and immersive interactions in VR.

PC-Based VR: PC-based VR systems, such as the Oculus Rift [www.oculus.com/rift-s] and HTC Vive [www.vive.com/us], require powerful computers with dedicated GPUs to run high-fidelity VR content.

- Stand-alone VR: Stand-alone VR headsets, like the Meta Quest [www.meta.com/quest], do not require a PC or console and are self-contained, making them more accessible and portable.
- Wireless VR: Wireless VR adapters eliminate the need for tethering the HMD to a computer, offering greater freedom of movement.
- Eye Tracking: Eye-tracking technology is being incorporated into VR HMDs for more realistic rendering, which reduces GPU load by rendering only the part of the scene the user is currently looking at.
- Haptic Gloves: Haptic devices provide tactile feedback to users, allowing them to feel the virtual or augmented environment, enhancing the sense of immersion. Haptic gloves are designed with sensors and actuators to provide realistic feedback to users' hands. This allows users to touch and manipulate virtual objects, making it crucial for applications like VR training simulations and virtual surgeries.
- Haptic Vests and Suits: These wearable haptic devices cover the user's body with multiple haptic points, enabling sensations like impact, pressure and vibration. They are used in gaming, training and medical simulations to simulate various physical interactions.
- Force Feedback Devices: Force feedback steering wheels and joysticks provide resistance and vibration during virtual driving and flying experiences, improving the sensation of control and interaction.
- Haptic Sound Systems: These systems use sound waves to create physical sensations. Ultrasonic haptic speakers can make users feel vibrations in the air, which is particularly useful for VR experiences.
- Bone Conduction Headphones: Bone conduction headphones transmit audio vibrations through the bones in the user's skull, allowing them to hear sounds while maintaining spatial awareness in AR applications.

1.4 Multimedia Software

Multimedia production relies heavily on a wide array of application software that facilitates content creation, editing and enhancement. Editing software is used to create and edit specific media components like image, graphics, audio, video, animation while authoring software is required to combine all edited media into a single integrated presentation, typically a runtime format either stand-alone like .EXE or a Web-based format like.SWF. As Web-based content creation developed, languages like HTML5, CSS, JavaScript as well as multimedia APIs and libraries were extensively used to create rich and interactive content that can be accessed by a global audience.

1.4.1 Editing Software

1.4.1.1 Text Editing

Text is a fundamental component of multimedia, and its effective integration into multimedia content requires specialized tools. These tools provide features for text editing, formatting and layout design. Major text editing tools include the following:

- Word Processors: Word processing software, such as Microsoft Word [microsoft.com/en-us/microsoft-365/word] and Google Docs [www.google.com/docs], offers features for creating, formatting and editing text. They are essential for developing scripts, articles and textual content in multimedia projects.
- Desktop Publishing Software: Desktop publishing software like Adobe InDesign [www.adobe.com/products/indesign] is employed for designing visually appealing documents and layouts. It is valuable for multimedia projects that involve brochures, magazines and print materials.

1.4.1.2 Image Editing

Images are a crucial component in multimedia production. Image editing software empowers creators to manipulate, enhance and optimize images for various applications. Prominent image editing tools include the following:

- Adobe Photoshop: Adobe Photoshop [www.adobe.com/products/photoshop] is a versatile image editing software that enables advanced image manipulation, retouching, and creative design. It is widely used in multimedia projects for image enhancement and compositing.
- GIMP: The GNU Image Manipulation Program (GIMP) [www.gimp.org] is an open-source alternative to Adobe Photoshop. It provides many of the same features for image editing and is favored by many multimedia professionals.

1.4.1.3 Audio Editing

Audio is integral to multimedia content, and dedicated software is used for audio recording, editing and production. These tools enable sound engineers, musicians and content creators to craft audio components. Major audio editing tools include the following:

Adobe Audition: Adobe Audition [www.adobe.com/products/audition] is a professional audio editing software that offers a comprehensive set of tools for recording, editing and mixing audio content, including features like waveform editing, spectral frequency display, and a robust multi-track environment. It provides advanced audio restoration capabilities, real-time effects and seamless integration with other Adobe Creative Cloud applications, making it a versatile choice for audio professionals and content creators.

Audacity: Audacity [www.audacityteam.org] is a free and open-source audio editing software known for its simplicity and versatility, offering features like multi-track editing, audio recording and precise waveform editing. It supports a wide range of audio formats, includes built-in effects and tools for noise reduction, and is accessible for users across different skill levels, making it a popular choice for various audio editing needs.

1.4.1.4 Video Editing

Video is a dynamic multimedia format, and video editing software is essential for creating, editing and refining video content. Leading video editing tools include the following:

Adobe Premiere Pro: Adobe Premiere Pro [www.adobe.com/products/premiere] is a professional video editing software with a range of powerful features, including a seamless workflow for editing, color grading, audio processing and visual effects in a single application. It offers an intuitive interface, advanced timeline editing, support for various video formats, and integration with other Adobe Creative Cloud applications, making it a popular choice for filmmakers, video editors and content creators.

Apple Final Cut Pro X: Final Cut Pro X [www.apple.com/final-cut-pro] is a professional video editing software known for its optimized performance on macOS, featuring a timeline for intuitive editing, advanced color grading tools, and a range of powerful audio editing capabilities. With a focus on speed and efficiency, it also offers features like multi-cam editing, 360-degree video editing and seamless integration with other Apple devices, making it a preferred choice for many filmmakers and video editors.

Avid Media Composer: Avid Media Composer [www.avid.com/media-composer] is a leading professional video editing software widely used in the film and television industry, offering robust tools for precise editing, color correction and audio post-production. Known for its reliable media management capabilities and collaborative workflow features, it provides a comprehensive solution for large-scale video projects.

Corel VideoStudio: Corel VideoStudio [www.videostudiopro.com] is designed for enthusiasts and beginners, offering an intuitive interface with drag-and-drop functionality, as well as a wide range of creative tools for editing, enhancing and sharing videos. With features like multi-camera editing, 3D title creation and built-in templates, it provides a versatile and accessible platform for individuals looking to create polished and engaging video content.

1.4.1.5 Graphics and Animation

2D and 3D graphics and animation tools enable creation, editing and publishing of 2D and 3D vector graphics and animation content. Usually incorporate a run-time software to render the graphics during playback.

Adobe Animate: Adobe Animate [www.adobe.com/products/animate.html] is a 2D animation tool known for its robust features in creating interactive and animated content for various platforms, including Web and mobile applications. With support for vector graphics, a timeline-based interface and the ability to publish content in multiple formats, Adobe Animate is widely used for designing engaging animations, interactive games and dynamic Web content.

Autodesk 3d Studio Max: Autodesk 3ds Max [www.autodesk.com/products/3ds-max] is a powerful 3D modeling and animation software widely utilized in the fields of architecture, gaming and visual effects production.

With a comprehensive set of modeling tools, advanced rendering capabilities and support for complex animations, it enables artists and designers to create realistic 3D content for various industries.

Autodesk Maya: Autodesk Maya [www.autodesk.com/products/maya] is a 3D animation and modeling software known for its versatile tools and capabilities, catering to industries such as film, television and game development. With advanced features like character rigging, dynamic simulations and a robust rendering engine, Maya empowers artists and animators to create highly realistic and visually stunning 3D content for a wide range of applications.

1.4.2 Authoring Software

Multimedia integration or authoring tools are essential components of the content creation process, enabling the seamless amalgamation of various multimedia elements such as text, images, audio, video and animation. These tools come in both stand-alone and Web-based formats, offering diverse capabilities to cater to the specific requirements of multimedia projects.

1.4.2.1 Stand-alone Authoring

Stand-alone multimedia integration software is installed on a local computer, providing a dedicated environment for content creators to compile, edit and synchronize multimedia components. The output file is usually published as an executable with its own run-time engine, for distribution on various platforms, including CD-ROMs, the Web using plug-ins and stand-alone applications for PCs. Some well-known stand-alone multimedia integration tools include the following.

Macromedia/Adobe Director: Adobe Director [www.adobe.com/support/director/downloads.html], initially developed by Macromedia and later acquired by Adobe, was a powerful multimedia authoring platform used for creating interactive multimedia applications. It allowed content creators to combine text, images, audio, video and animations into cohesive presentations or games. Its integrated scripting language Lingo was used for creating interactive and engaging multimedia content. However, it has been succeeded by Adobe Animate in more recent years. In 2017, Adobe officially announced the discontinuation of Adobe Director.

Macromedia Authorware/Adobe Captivate: Macromedia Authorware, also developed by Macromedia, now Adobe Captivate [www.adobe.com/products/captivate], is renowned for its capability to create e-learning and training applications. It employs a flowchart-based visual authoring environment that enabled content creators to integrate multimedia elements, quizzes and interactive simulations. Macromedia Authorware was particularly valuable for educational multimedia. Macromedia Authorware was officially discontinued by Adobe in the year 2007.

Adobe Flash: Adobe Flash [www.adobe.com/products/flashruntimes.html], while no longer actively developed by Adobe, played an important role in multimedia content creation, especially on the Web. Flash enabled the creation of interactive animations, games and multimedia-rich websites. Its integrated scripting language called ActionScript was used to add interactivity and animations to content. With the rise of HTML5 and Web standards, Adobe phased out Flash, and Web browsers discontinued support, rendering it obsolete.

Asymetrix Toolbook: Asymetrix Toolbook was a multimedia authoring tool and development environment that gained popularity in the 1990s and early 2000s. Toolbook, created by Asymetrix, now Sum Total Systems [www.sumtotalsystems.com], allowed developers to create interactive multimedia content for various purposes, including e-learning courses, multimedia presentations, simulations and interactive applications. Toolbook featured a visual, drag-and-drop interface that made it accessible to users with varying levels of technical expertise. Its integrated scripting language called OpenScript was used for adding interactivity and functionality to multimedia projects. OpenScript allowed developers to create custom behaviors and interactions within their content.

1.4.2.2 Web-Based Authoring

As Web technologies evolved, the demand for CD-ROM-based multimedia content also diminished. Director's strengths in creating interactive CD-ROMs and stand-alone applications were less relevant in a world dominated by Web-based content, mobile apps and responsive design. This shift in Web standards, combined with the rise of HTML5 and JavaScript, made it increasingly challenging for these tools to remain relevant. Web-based multimedia authoring software is accessible through Web browsers and allows collaboration from different locations. These tools are often preferred for their ease of access and collaborative features. Notable Web-based multimedia authoring tools include the following:

Canva: Canva [www.canva.com] is a user-friendly graphic design platform that offers a wide range of features for creating visually appealing content. With an extensive library of templates, customizable design elements and a drag-and-drop interface, it enables users to effortlessly design social media graphics, presentations, posters and more.

Prezi: Prezi [www.prezi.com] is a dynamic presentation software that distinguishes itself from traditional slideshows by offering a zoomable canvas, allowing users to create non-linear and engaging presentations. With features like motion paths, interactive elements and the ability to present ideas in a visually compelling way, it promotes a more dynamic and memorable presentation experience.

Crello: Crello [www.appsumo.com/products/crello] is a Web-based graphic design tool that allows users to create static and animated designs for social media, advertising and presentations. It provides access to a vast library of templates and design elements.

Kapwing: Kapwing [www.kapwing.com] is a Web-based video editing platform that simplifies video creation and editing. It offers a range of tools for trimming, adding text, memes, collages and creating engaging video content for social media.

Articulate Storyline: Articulate Storyline [www.articulate.com/360/storyline] is a robust e-learning authoring tool known for its powerful features in creating interactive and engaging online courses. Its major features include a user-friendly interface, drag-and-drop functionality and a vast library of customizable templates for building multimedia-rich content, with support for branching scenarios, quizzes and responsive design.

H5P: H5P [www.h5p.org] is an open-source platform for creating and sharing interactive HTML5 content. It's a versatile tool for building interactive multimedia content, including presentations, quizzes and games.

Tumult Hype: Tumult Hype [www.tumult.com/hype] is known for its focus on creating interactive and engaging HTML5 content. It provides a user-friendly interface for designing animations, presentations and interactive Web content without requiring extensive coding knowledge.

Vyond (formerly GoAnimate): Vyond [www.vyond.com] enables users to create animated videos without the need for extensive coding skills. Its user-friendly interface allows individuals, educators and businesses to produce professional-looking animations using a drag-and-drop approach.

Genially: Genially [www.genial.ly] is a versatile online tool designed for creating interactive and visually appealing content such as presentations, info-graphics and educational materials. It offers a variety of templates across different categories, making it convenient for users to start their projects with pre-designed structures.

Adobe Spark: Adobe Spark [express.adobe.com/page/9bPqZ] offers a suite of Web-based tools for creating social media graphics, Web pages and short videos. It simplifies graphic design for users without extensive design skills, offering an intuitive platform to create visually appealing graphics, Web pages and short videos.

Ceros: Ceros [www.ceros.com] is a cloud-based interactive content creation platform, supporting real-time collaboration, allowing teams to work together seamlessly on projects. It is designed for marketers and content creators to build interactive Web experiences, including info-graphics, microsites and digital magazines.

Wix: Wix [www.wix.com] is a website builder that offers a range of multimedia and interactive features. It enables users to create visually appealing websites with multimedia elements, including videos, animations and interactive forms.

1.4.3 Multimedia APIs and Libraries

Multimedia is a rich and diverse field encompassing a wide range of content types, including images, audio, video, and interactive applications. As Web-based multimedia content creation gained momentum, developers relied on multimedia Application Programming Interfaces (APIs) and libraries, to provide the building blocks necessary to harness the power of Web-based multimedia. Developers could integrate multimedia APIs and libraries into their applications by including the necessary libraries, linking them to their code and leveraging their functions and classes. Programming languages such as C++, Python, Java and JavaScript can also interact with multimedia APIs. Specific libraries may have language bindings or wrappers to facilitate integration. These APIs and libraries provide methods and functions to perform operations like resizing images, applying filters, converting media formats, extracting metadata and more. They also enable multimedia streaming, live broadcasting and playback functionality, making them essential for applications like video conferencing, online media streaming and gaming.

Image Processing: Libraries like OpenCV [www.opencv.ai] and Pillow (PIL) [www.python-pillow.org] offer extensive image processing capabilities, enabling tasks such as image editing, manipulation, filtering and computer vision applications. OpenCV (Open Source Computer Vision) library is a powerful open-source library designed for computer vision and machine learning applications with a comprehensive set of tools and functions that allow developers to build and implement computer vision algorithms, image and video processing, and object detection. Pillow, a Python Imaging Library (PIL) fork, is a comprehensive image processing library, designed to be cross-platform and works seamlessly on different operating systems, including Windows, macOS and Linux.

Audio Processing: Libraries like Libsndfile [libsndfile.github.io/libsndfile] and PyDub [www.pypi.org/project/pydub] provide tools for audio file format conversion, processing, editing and streaming. Libsndfile is a C library that provides functionalities for reading and writing audio files in various formats, and supports multi-channel audio, allowing developers to work with stereo, surround sound, and other multi-channel configurations. PyDub is a Python library that simplifies audio processing tasks by providing a high-level interface for common audio operations, allowing users to manipulate audio files, including tasks such as cutting, concatenating and applying various effects, all with a straightforward syntax.

Video Processing: FFmpeg [www.ffmpeg.org] is a powerful multimedia framework that encompasses a vast set of tools and libraries for handling audio and video processing tasks. Its major feature lies in its ability to decode, encode, transcode, multiplex, de-multiplex, stream and filter multimedia content.

Graphics and 2D/3D Rendering: APIs like OpenGL [www.opengl.org], Vulkan [www.vulkan.org] and DirectX [learn.microsoft.com/en-us/windows/win32/directx] are used for creating and rendering graphics and 3D models in interactive applications, including games and simulations. OpenGL is an open-source graphics API that facilitates cross-platform rendering of 2D and 3D graphics by taking advantage of hardware acceleration and utilizing the capabilities of graphics processing units (GPUs). DirectX is a collection of APIs developed by Microsoft, primarily designed for multimedia and gaming applications on the Windows platform and includes components like Direct3D, DirectSound and DirectShow for graphics, audio and playback systems. Vulkan is a low-level, cross-platform graphics API designed for high-performance graphics rendering across different operating systems with explicit control over the GPU, enabling more efficient use of hardware resources and allowing for advanced graphics rendering techniques.

Multimedia Playback: Libraries such as VLC [www.videolan.org/vlc] and GStreamer [gstreamer.freedesktop.org] allow developers to integrate multimedia playback into their applications. VLC (VideoLAN Client) is a free and open-source versatile and widely used media player known for its cross-platform compatibility, supporting various operating systems including Windows, macOS, Linux, Android and iOS and its ability to play a vast array of multimedia file formats. GStreamer is an open-source multimedia framework designed for building multimedia applications across different platforms supporting a modular architecture, and offering a pipeline-based model that enables developers to create complex multimedia workflows.

User Interface Development: Frameworks like Qt [www.qt.io], GTK [www.gtk.org] and JavaFX [openjfx.io] enable the creation of multimedia-rich user interfaces with features like multimedia playback, image rendering and animations. Qt is a powerful cross-platform application development framework that enables developers to create applications that run seamlessly on various operating systems, including Windows, macOS, Linux, Android and iOS and provides an extensive set of tools and libraries for creating graphical user interfaces (GUIs) with a wide range of widgets and styles. GTK (GIMP Toolkit) is a cross-platform widget toolkit used for creating graphical user interfaces (GUIs) and provides an extensive set of widgets and tools for creating modern and responsive user interfaces. JavaFX is a Java library for creating rich GUIs with a diverse set of built-in UI components, including a comprehensive collection of customizable UI elements, such as charts, tables and multimedia support.

Web-Based Multimedia: Web technologies like HTML5, CSS3, and JavaScript offer powerful capabilities for embedding multimedia content within Web applications, allowing for audio and video playback, animations, and interactive multimedia elements. HTML5 introduces native support for multimedia elements such as audio and video, eliminating the need for third-party plugins like Flash, which enables developers to embed and control multimedia content directly within Web pages. CSS3 introduces a wide range of new styling features, including gradients, shadows, and flexible box layouts, providing developers with more versatile and expressive tools for designing modern Web layouts. JavaScript14 introduces new set of functions for manipulating, searching and sorting arrays.

1.5 Multimedia Standards

Multimedia standards are essential in the world of digital content creation, storage, transmission and playback. They define the technical specifications, formats and protocols that ensure consistency and interoperability across various multimedia devices and platforms. These standards help in the seamless exchange and playback of multimedia content, as well as collaborative ventures in content creation.

CODEC Standards: Multimedia CODECs (coder-decoder) play a crucial role in encoding and decoding image, audio and video data. Popular codec standards include JPEG2000 [jpeg.org/jpeg2000] for image, H.264 [itu.int/rec/T-REC-H.264] for video and AAC [mpeg.chiariglione.org/standards/mpeg-4/audio] for audio, while emerging standards like H.265 (or HEVC – High Efficiency Video Coding) [itu.int/rec/T-REC-H.265] and AV1 [aomedia.org/av1-features] aim to improve compression efficiency.

Streaming Protocols: Standards like HTTP Live Streaming (HLS) [developer.apple.com/streaming], Dynamic Adaptive Streaming over HTTP (DASH) [mpeg.chiariglione.org/standards/mpeg-dash] and Real-Time Messaging Protocol (RTMP) [helpx.adobe.com/adobe-media-server/dev/stream-on-demand-media-rtmp.html] govern the delivery of multimedia content over the internet, ensuring adaptive streaming for various bandwidth conditions.

Image Standards: Image formats like JPEG [jpeg.org], PNG [libpng.org/pub/png], GIF [w3.org/Graphics/GIF/spec-gif89a.txt] and WebP [developers.google.com/speed/webp] are standardized for still images, offering varying levels of compression, quality and support for transparency.

Audio Standards: Audio standards include formats like MP3 [mpeg.chiariglione.org/standards/mpeg-1/audio] [iso.org/standard/22412.html], WAV [learn.microsoft.com/en-us/windows/win32/multimedia/waveform-audio] and FLAC [xiph.org/flac] for audio encoding and playback. The development of spatial audio standards like Dolby Atmos [dolby.com/technologies/dolby-atmos] and DTS:X [dts.com/dtsx] has transformed the audio experience.

Video Container Formats: Container formats, such as MP4 [mpeg.chiariglione.org/standards/mpeg-4/video], MKV [matroska.org], and AVI [learn.microsoft.com/en-us/windows/win32/directshow/avi-file-format], wrap multimedia content, including video, audio, subtitles and metadata, into a single file. These formats are standardized to ensure compatibility across various devices and players.

Metadata Standards: Metadata standards like Exif [cipa.jp/e/std/std-sec] for images and ID3 [mpgedit.org/mpgedit/mpeg_format/mpeghdr.htm#mpegtag] for audio allow multimedia content to carry information such as date, location, artist, album and copyright details.

Color Management Standards: Standards like sRGB [webstore.iec.ch/publication/6169], AdobeRGB [color.org/chardata/rgb/adobergb.xalter] and ProPhoto/ROMM RGB (reference output medium metric) [iso.org/standard/56591.html] define color spaces, ensuring consistent color reproduction across devices and platforms.

Captioning and Subtitle Standards: Formats like WebVTT (Web Video Text Track) [w3.org/TR/webvtt] and TTML (Timed Text Markup Language) [w3.org/TR/ttml] are used to create captions and subtitles for multimedia content, making it accessible to a broader audience. WebVTT is a standard designed for adding captions, subtitles and other text tracks to HTML5 videos, TTML is a standard for representing timed text, such as captions and subtitles, in online media content.

Medical imaging standard: The DICOM (Digital Imaging and Communications in Medicine) standard [dicomstandard.org] provides a foundational framework for medical imaging and healthcare informatics by defining a structure and format of medical images as also the communication and networking protocols required for seamless data exchange within healthcare facilities.

1.5.1 JPEG Overview

JPEG (Joint Photographic Experts Group) [jpeg.org] is a standard for lossy compression for digital photographic images.

1.5.1.1 JPEG 1

The original JPEG specification, called JPEG 1 [jpeg.org/jpeg], published in 1992, implements processes from various earlier research papers and patents cited by the CCITT (now ITU-T) and Joint Photographic Experts Group. The standard specifies the CODEC (coder/decoder) used to compress and decompress the image while the Exif and JFIF standards define the file format. The degree of compression can be adjusted by selecting a parameter for controlling image quality, but typically it achieves a 10:1 compression ratio with acceptable quality. The MIME type for JPEG is "image/jpeg" and a filename extension of ".jpg" or ".jpeg". The JPEG committee has as one of its explicit goals that their standards, especially their baseline methods, be implementable without payment of license fees. The original JPEG standard was adopted by ITU-T as Recommendation T.81 and by ISO as standard ISO/IEC 10918-1 [www.iso.org/standard/18902.html]. The standard consists of the following parts:

- Part 1: ISO/IEC 10918-1:1994 – Requirements and guidelines
- Part 2: ISO/IEC 10918-2:1995 – Compliance testing
- Part 3: ISO/IEC 10918-3:1997 –Extensions
- Part 4: ISO/IEC 10918-4:1999 – Registration authorities
- Part 5: ISO/IEC 10918-5:2013 – File Interchange Format
- Part 6: ISO/IEC 10918-6:2013 –Application to Printing Systems
- Part 7: ISO/IEC 10918-7:2023 –Reference Software

1.5.1.2 JPEG 2000

JPEG 2000 [jpeg.org/jpeg2000], formally ISO/IEC 15444 [www.iso.org/standard/78321.html], offered a number of benefits over the original JPEG format, namely, multiple resolutions of the same image, progressive decoding, choice of lossy or lossless modes, error resilience, flexible file format, HDR support, support for transparency and alpha channels. It consists of the following parts:

- Part 1: ISO/IEC 15444-1:2000 – Core coding system
- Part 2: ISO/IEC 15444-2:2004 – Extensions
- Part 3: ISO/IEC 15444-3:2002 – Motion JPEG 2000 (MJ2 or MJP2)
- Part 4: ISO/IEC 15444-4:2002 – Conformance testing
- Part 5: ISO/IEC 15444-5:2003 – Reference Software
- Part 6: ISO/IEC 15444-6:2003 – Compound image file format
- Part 7: abandoned
- Part 8: ISO/IEC 15444-8:2007 – Secure JPEG 2000
- Part 9: ISO/IEC 15444-9:2005 – Interactivity tools, APIs and protocols
- Part 10: ISO/IEC 15444-10:2008 – Extensions for three dimensional data (JP3D)
- Part 11: ISO/IEC 15444-11:2007 – Wireless (JPWL)

- Part 12: ISO/IEC 15444-12:2004 –Base media file format
- Part 13: ISO/IEC 15444-13:2008 –Entry level encoder
- Part 14: ISO/IEC 15444-14:2013 – XML representation (JPXML)
- Part 15: ISO/IEC 15444-15:2019 – High throughput format
- Part 16: ISO/IEC 15444-16:2019 – High efficiency file format

1.5.1.3 JPEG AI

JPEG AI [jpeg.org/jpegai], formally ISO/IEC CD 6048 [www.iso.org/standard/81984.html], is designed for creation of a learning-based image coding standard offering a single-stream, compact compressed domain representation, targeting both human visualization, with significant compression efficiency improvement over image coding standards in common use at equivalent subjective quality, and effective performance for image processing and computer vision tasks, with the goal of supporting a royalty-free baseline. It consists of the following parts:

- Part 1: ISO/IEC 6048-1 – Core coding system
- Part 2: ISO/IEC 6048-2 – Profiling
- Part 3: ISO/IEC 6048-3 – Reference software
- Part 4: ISO/IEC 6048-4 – Conformance

1.5.1.4 JPEG AIC

JPEG Advanced Image Coding and Evaluations [jpeg.org/aic], formally ISO/IEC 29170 [www.iso.org/standard/66094.html], is designed to locate and evaluate new scientific developments and advancements in image coding research. Relevant topics include new compression methodologies and quality evaluation methodologies and procedures. It consists of the following parts:

- Part 1: ISO/IEC 29170-1 – Guidelines for image coding system evaluation
- Part 2: ISO/IEC 29170-2 – Evaluation procedure for nearly lossless coding

1.5.1.5 JPEG Pleno

JPEG Pleno [jpeg.org/jpegpleno], formally ISO/IEC 21794 [www.iso.org/standard/74531.html], aims to provide a standard framework for representing new imaging modalities, such as texture-plus-depth, light field, point cloud, and holographic imaging. Such imaging should be understood as light representations inspired by the plenoptic function, which describe intensity of light rays as functions of visual angle, wavelength, time and visual angle. It consists of the following parts:

- Part 1: ISO/IEC 21794-1 – Framework
- Part 2: ISO/IEC 21794-2 – Light field coding
- Part 3: ISO/IEC 21794-3 – Conformance testing
- Part 4: ISO/IEC 21794-4 – Reference software

1.5.1.6 JPEG Systems

JPEG Systems [jpeg.org/jpegsystems], formally ISO/IEC 19566 [www.iso.org/standard/81643.html], is a suite of standards developed by the JPEG Committee that extend JPEG coding standards with wider system aspects and extensions. It describes file formats, transport mechanism, adds metadata and additional functionalities that are necessary to support applications beyond purely coded image data. It consists of the following parts:

- Part 1: ISO/IEC 19566-1 – Packaging of information using codestreams and file formats
- Part 2: ISO/IEC 19566-2 – Transport mechanisms and packaging
- Part 4: ISO/IEC 19566-4 – Privacy and security
- Part 5: ISO/IEC 19566-5 – JPEG universal metadata box format (JUMBF)
- Part 6: ISO/IEC 19566-6 – JPEG 360
- Part 7: ISO/IEC 19566-7 – JPEG linked media format (JLINK)
- Part 8: ISO/IEC 19566-8 – JPEG Snack

1.5.1.7 JPEG Trust

JPEG Trust [jpeg.org/jpegtrust] defines a framework for establishing trust in media. It addresses aspects of authenticity, provenance and integrity through secure and reliable annotation of media assets throughout their life cycle to address problems related to AI-based synthetic media such as deepfakes. The new standard is expected to be published in 2024.

1.5.1.8 JPEG XL

JPEG XL [jpeg.org/jpegxl], formally ISO/IEC 18181 [www.iso.org/standard/77977.html], is designed to meet the needs of image delivery on the Web and professional photography. It supports wide color gamut as well as high dynamic range and high bit depth images. JPEG XL further includes features such as animation, alpha channels, layers, thumbnails, lossless and progressive coding to support a wide range of use cases, including, but not limited to, photo galleries, e-commerce, social media, user interfaces and cloud storage. To enable novel applications, it also adds support for 360 degree images, image bursts, large panoramas/mosaics and printing. It consists of the following parts:

- Part 1: ISO/IEC 18181-1:2022 – Core coding system
- Part 2: ISO/IEC 18181-2:2021 – File format
- Part 3: ISO/IEC 18181-3:2022 – Conformance testing
- Part 4: ISO/IEC 18181-4:2022 – Reference software

1.5.1.9 JPEG XR

JPEG XR [jpeg.org/jpegxr], formally ISO/IEC 29199 [www.iso.org/standard/79414.html], provides a practical coding technology for a broad range of applications with excellent compression capability and important additional functionalities. The primary intended application of JPEG XR is the representation of continuous-tone still images such as photographic images. It consists of the following parts:

- Part 1: ISO/IEC 29199-1 – System architecture
- Part 2: ISO/IEC 29199-2 – Image coding specification
- Part 3: ISO/IEC 29199-3 – Motion JPEG XR
- Part 4: ISO/IEC 29199-4 – Conformance testing
- Part 5: ISO/IEC 29199-5 – Reference software

1.5.1.10 JPEG XS

JPEG XS [jpeg.org/jpegxs], formally ISO/IEC 21122 [www.iso.org/standard/81551.html], is an extension for visually lossless, low-latency and lightweight files, designed for applications like virtual reality, drones and self-driving cars. It consists of the following parts:

- Part 1: ISO/IEC 21122-1 – Core coding system
- Part 2: ISO/IEC 21122-2 – Profiles and buffer models
- Part 3: ISO/IEC 21122-3 – Transport and container formats
- Part 4: ISO/IEC 21122-4 – Conformance testing
- Part 5: ISO/IEC 21122-5 – Reference software

1.5.1.11 JPEG XT

JPEG XT [jpeg.org/jpegxt], formally ISO/IEC 18477 [www.iso.org/standard/79103.html], is a backward compatible extension to the legacy JPEG standard with additional support for higher integer bit depths, HDR imaging, lossless coding, alpha channel coding and extensible file format. It consists of the following parts:

- Part 1: ISO/IEC 18477-1 – Core coding system
- Part 2: ISO/IEC 18477-2 – HDR (High dynamic range) imaging
- Part 3: ISO/IEC 18477-3 – File format
- Part 4: ISO/IEC 18477-4 – Conformance testing
- Part 5: ISO/IEC 18477-5 – Reference software
- Part 6: ISO/IEC 18477-6 – IDR Integer coding
- Part 7: ISO/IEC 18477-7 – HDR Floating-Point Coding
- Part 8: ISO/IEC 18477-8 – Lossless and Near-lossless Coding
- Part 9: ISO/IEC 18477-9 – Alpha channel Coding

1.5.1.12 JPEG LS

JPEG LS [jpeg.org/jpegls], formally ISO/IEC 14495 [www.iso.org/standard/22397.html], was defined to address the need for effective lossless and near-lossless compression of continuous-tone still images. JPEG LS is especially suited for low-complexity hardware implementations of very moderate complexity, while at the same time providing state-of-the-art lossless compression performance. It consists of the following parts:

- Part 1: ISO/IEC 14495-1 – Baseline
- Part 2: ISO/IEC 14495-2 – Extensions

1.5.1.13 JPSearch

JPSearch [jpeg.org/jpsearch], formally ISO/IEC 24800 [www.iso.org/standard/78323.html], is designed to address interoperability in image search and retrieval systems. For this purpose, JPSearch puts forward an abstract image search and retrieval framework. Interfaces and protocols for data exchange between the components of this architecture are standardized. It consists of the following parts:

- Part 1: ISO/IEC 24800-1 – System framework and components
- Part 2: ISO/IEC 24800-2 – Registration, identification and management of schema and ontology

- Part 3: ISO/IEC 24800-3 – Query format
- Part 4: ISO/IEC 24800-4 – File format
- Part 5: ISO/IEC 24800-5 – Data interchange format
- Part 6: ISO/IEC 24800-6 – Reference software

1.5.1.14 JBIG

JBIG1 [jpeg.org/jbig] is an early lossless image compression standard for binary images released by the Joint Bi-level Image Experts Group, in 1993 as standard ISO/IEC 11544 [www.iso.org/standard/19498.html], and widely used in fax machines, Later in 2000, an improved version JBIG2 with better coding efficiency was released as standard ISO/IEC 14492 [www.iso.org/standard/75396.html]. It consists of the following parts:

- ISO/IEC 11544:1993 – Coded representation of picture and audio information
- ISO/IEC 14492:2019 – Lossy/lossless coding of bi-level images

1.5.2 MPEG Overview

The Motion Pictures Expert Group (MPEG) [mpeg.chiariglione.org] is a working group under ISO/IEC (International Standards Organization/International Electrotechnical Commission) set up to formulate a set of standards relating to a range of multimedia application involving audio and video primarily. Since its first meeting in 1988, MPEG has grown to include approximately 350 members from the industry and academia. MPEG's official designation is ISO/IEC JCT1/SC29 WG11 [www.iso.org/committee/45316.html]. This section contains an overview of the various standards developed.

1.5.2.1 MPEG-1

The International Standards Organization (ISO) and the International Electro-technical Commission (IEC) formed the Moving Pictures Experts Group (MPEG) in 1988 to devise compression techniques for audio and video. This group has developed a number of highly successful standards. It first devised the ISO/IEC International Standard 11172 for reduced data-rate coding of video and audio signals. The standard was finalized in November 1992 and is commonly known as MPEG-1 [mpeg.chiariglione.org/standards/mpeg-1]. It was later used as a standard for Video CD and includes the popular MP3 audio format. It is defined in a series of documents all subsets of ISO/IEC Recommendation 11172. It consists of the following parts:

- Part 1: ISO/IEC 11172-1:1993 – Systems
- Part 2: ISO/IEC 11172-2:1993 – Video
- Part 3: ISO/IEC 11172-3:1993 – Audio
- Part 4: ISO/IEC 11172-4:1995 – Compliance Testing
- Part 5: ISO/IEC 11172-5:1998 – Software Simulation

1.5.2.2 MPEG-2

MPEG-2 [mpeg.chiariglione.org/standards/mpeg-2], formally ISO/IEC 13818, is a video and audio storage and transmission standards for broadcast quality television. It is used for digital TV transmission, digital satellite TV services, digital cable TV signals and for DVD applications. It is defined in a series of documents which are all subsets of ISO Recommendations 13818. The standard currently has nine parts: the first three have reached international standard status, while the others are at different levels of completion; one has been withdrawn. It consists of the following parts:

- Part 1: ISO/IEC 13818-1:2000 – Systems
- Part 2: ISO/IEC 13818-2:2000 – Video
- Part 3: ISO/IEC 13818-3:1998 – Audio
- Part 4: ISO/IEC 13818-4:1998 – Conformance Testing
- Part 5: ISO/IEC 13818-5:1997 – Software Simulation S
- Part 6: ISO/IEC 13818-6:1998 – Extensions for DSM-CC (Digital Storage Media Command and Control)
- Part 7: ISO/IEC 13818-7:1997 – Advanced Audio Coding AAC
- Part 8: ISO/IEC 13818-8:1997 – abandoned
- Part 9: ISO/IEC 13818-9:1996 – Extensions for Real-time Interface for System Decoders
- Part 10: ISO/IEC 13818-10:1999 – Conformance Extensions for DSM-CC

1.5.2.3 MPEG-4

MPEG-4 [mpeg.chiariglione.org/standards/mpeg-4], formally ISO/IEC 14496, introduced in late 1998, is the designation for a group of standards for coding of audio visual objects and related technology agreed upon by the ISO/IEC Moving Picture Experts Group (MPEG). The primary uses for the MPEG-4 standard are Web (streaming media) and CD distribution, conversational (videophone), and broadcast television. MPEG-4 absorbs many of the features of MPEG-1 and MPEG-2 and other related standards, adding new features such as (extended) VRML support for 3D rendering, object-oriented composite files (including audio, video, and VRML (Virtual Reality Modeling Language objects), support for externally specified Digital Rights Management and various types of interactivity. Most of the features included in MPEG-4 are left to individual developers to decide whether to implement them. This means that there are probably no complete implementations of the entire MPEG-4 set of standards. To deal with this, the standard includes the concept of profiles and levels, allowing a specific set of capabilities to be defined in a manner appropriate for a subset of applications. It consists of the following parts:

- Part 1: ISO/IEC 14496-1 – Systems
- Part 2: ISO/IEC 14496-2 – Visual
- Part 3: ISO/IEC 14496-3 – Audio
- Part 4: ISO/IEC 14496-4 – Conformance
- Part 5: ISO/IEC 14496-5 – Reference Software
- Part 6: ISO/IEC 14496-6 – Delivery Multimedia Integration Framework (DMIF)
- Part 7: ISO/IEC 14496-7 – Optimized Reference Software
- Part 8: ISO/IEC 14496-8 – Carriage on IP Networks
- Part 9: ISO/IEC 14496-9 – Reference Hardware
- Part 10: ISO/IEC 14496-10 – Advanced Video Coding (AVC)
- Part 11: ISO/IEC 14496-11 – Scene Description and Application Engine
- Part 12: ISO/IEC 14496-12 – ISO Base Media File Format
- Part 13: ISO/IEC 14496-13 – Intellectual Property Management and Protection (IPMP) Extensions
- Part 14: ISO/IEC 14496-14 – MPEG-4 File Format
- Part 15: ISO/IEC 14496-15 – Advanced Video Coding (AVC) File Format
- Part 16: ISO/IEC 14496-16 – Animation Framework eXtension (AFX)
- Part 17: ISO/IEC 14496-17 – Timed Text Subtitle Format
- Part 18: ISO/IEC 14496-18 – Font Compression and Streaming.
- Part 19: ISO/IEC 14496-19 – Synthesized Texture Stream
- Part 20: ISO/IEC 14496-20 – Lightweight Application Scene Representation (LASeR)
- Part 21: ISO/IEC 14496-21 – MPEG-J Graphics Framework eXtensions (GFX)
- Part 22: ISO/IEC 14496-22 – Open Font Format
- Part 23: ISO/IEC 14496-23 – Symbolic Music Representation (SMR)
- Part 24: ISO/IEC 14496-24 – Audio and Systems Interaction
- Part 25: ISO/IEC 14496-25 – 3D Graphics Compression Format
- Part 26: ISO/IEC 14496-26 – Audio Conformance
- Part 27: (ISO/IEC 14496-27) 3D Graphics Conformance

1.5.2.4 MPEG-5

MPEG-5 [mpeg.chiariglione.org/standards/mpeg-5], formally ISO/IEC 23094, is an evolution of MPEG-4 aiming to improve compression efficiency and support emerging technologies like 3D video and ultra-high definition. It introduces new tools for coding efficiency and provides enhanced features for video coding. MPEG-5 builds upon the foundations of MPEG-4 but incorporates advancements to deliver better video coding performance and accommodate evolving industry needs. It consists of the following parts:

- Part 1: ISO/IEC 23094-1 – Essential Video Coding (EVC)
- Part 2: ISO/IEC 23094-2 – Low Complexity Enhancement Video Coding (LCEVC)
- Part 3: ISO/IEC 23094-3 – Attribute value coding for ISO base file format

1.5.2.5 MPEG-7

MPEG-7 [mpeg.chiariglione.org/standards/mpeg-7], formally ISO/IEC 15938, is a suite of standards for description and search of audio, visual and multimedia content, referred to as Multimedia content description interface. It consists of the following parts:

- Part 1: ISO/IEC 15938-1 – Systems
- Part 2: ISO/IEC 15938-2 – Description Definition Language
- Part 3: ISO/IEC 15938-3 – Visual
- Part 4: ISO/IEC 15938-4 – Audio

- Part 5: ISO/IEC 15938-5 – Multimedia Description Schemes
- Part 6: ISO/IEC 15938-6 – Reference Software
- Part 7: ISO/IEC 15938-7 – Conformance Testing
- Part 8: ISO/IEC 15938-8 – Extraction and use of MPEG-7 descriptions
- Part 9: ISO/IEC 15938-9 – Profiles and Levels
- Part 10: ISO/IEC 15938-10 – Schema Definitions
- Part 11: ISO/IEC 15938-11 – Profile Schemas
- Part 12: ISO/IEC 15938-12 – Query format
- Part 13: ISO/IEC 15938-13 – Compact Descriptors for Visual Search (CDVS)
- Part 14: ISO/IEC 15938-14 – Reference software, conformance and usage guidelines for CDVS
- Part 15: ISO/IEC 15938-15 – Compact descriptors for video analysis (CDVA)
- Part 16: ISO/IEC 15938-16 – Conformance and Reference Software for CDVA
- Part 17: ISO/IEC 15938-17 – Compression of neural networks for multimedia content description and analysis

1.5.2.6 MPEG-21

MPEG-21 [mpeg.chiariglione.org/standards/mpeg-21], formally ISO/IEC 21000, is a suite of standard that define a normative open framework for end-to-end multimedia creation, delivery and consumption that provides content creators, producers, distributors and service providers with equal opportunities in the MPEG-21 enabled open market, and also be to the benefit of the content consumers providing them access to a large variety of content in an interoperable manner. It consists of the following parts:

- Part 1: ISO/IEC 21000-1 – Vision, Technologies and Strategy
- Part 2: ISO/IEC 21000-2 – Digital Item Declaration
- Part 3: ISO/IEC 21000-3 – Digital Item Identification
- Part 4: ISO/IEC 21000-4 – IPMP (Intellectual Property Management and Protection) Components
- Part 5: ISO/IEC 21000-5 – Rights Expression Language
- Part 6: ISO/IEC 21000-6 – Rights Data Dictionary
- Part 7: ISO/IEC 21000-7 – Digital Item Adaptation
- Part 8: ISO/IEC 21000-8 – Reference Software
- Part 9: ISO/IEC 21000-9 – File Format
- Part 10: ISO/IEC 21000-10 – Digital Item Processing
- Part 11: ISO/IEC 21000-11 – Evaluation Tools for Persistent Association
- Part 12: ISO/IEC 21000-12 – Test Bed for MPEG-21 Resource Delivery
- Part 13: ISO/IEC 21000-13 – void
- Part 14: ISO/IEC 21000-14 – Conformance
- Part 15: ISO/IEC 21000-15 – Event reporting
- Part 16: ISO/IEC 21000-16 – Binary format
- Part 17: ISO/IEC 21000-17 – Fragment Identification
- Part 18: ISO/IEC 21000-18 – Digital Item Streaming
- Part 19: ISO/IEC 21000-19 – White paper on MPEG IPR Ontologies
- Part 20: ISO/IEC 21000-20 – Contract Expression Language
- Part 21: ISO/IEC 21000-21 – Media Contract Ontology
- Part 22: ISO/IEC 21000-22 – User Description

1.5.2.7 MPEG-A

MPEG-A [mpeg.chiariglione.org/standards/mpeg-a], formally ISO/IEC 23000, is a group of standards for composing MPEG systems referred to as Multimedia Application Format. It consists of the following parts:

- Part 1: ISO/IEC 23000-1 – Purpose for multimedia application formats
- Part 2: ISO/IEC 23000-2 – MPEG music player application format
- Part 3: ISO/IEC 23000-3 – MPEG photo player application format
- Part 4: ISO/IEC 23000-4 – Musical slide show application format
- Part 5: ISO/IEC 23000-5 – Media streaming application format
- Part 6: ISO/IEC 23000-6 – Professional archival application format
- Part 7: ISO/IEC 23000-7 – Open access application format

- Part 8: ISO/IEC 23000-8 – Portable video application format
- Part 9: ISO/IEC 23000-9 – Digital Multimedia Broadcasting application format
- Part 10: ISO/IEC 23000-10 – Surveillance application format
- Part 11: ISO/IEC 23000-11 – Stereoscopic video application format
- Part 12: ISO/IEC 23000-12 – Interactive music application format
- Part 13: ISO/IEC 23000-13 – Augmented reality application format
- Part 14: ISO/IEC 23000-14 – Closed (was Mixed and Augmented Reality Reference Model)
- Part 15: ISO/IEC 23000-15 – Multimedia Preservation Application Format
- Part 16: ISO/IEC 23000-16 – Publish/Subscribe Application Format
- Part 17: ISO/IEC 23000-17 – Multisensory Effects Application Format
- Part 18: ISO/IEC 23000-18 – Media Linking Application Format
- Part 19: ISO/IEC 23000-19 – Common Media Application Format
- Part 20: ISO/IEC 23000-20 – Visual Identity Application Format

1.5.2.8 MPEG-B

MPEG-B [mpeg.chiariglione.org/standards/mpeg-b], formally ISO/IEC 23001, is a suite of standards for systems technologies that do not fall in other well-established MPEG standards. It consists of the following parts:

- Part 1: ISO/IEC 23001-1 – Binary MPEG format for XML
- Part 2: ISO/IEC 23002-2 – Fragment Request Unit
- Part 3: ISO/IEC 23003-3 – XML Representation of IPMP-X messages
- Part 4: ISO/IEC 23001-4 – Codec Configuration Representation
- Part 5: ISO/IEC 23001-5 – Bitstream Syntax Description Language
- Part 6: ISO/IEC 23001-6 – void (to MPEG-DASH part 1)
- Part 7: ISO/IEC 23001-7 – Common Encryption for ISO Base Media File Format Files
- Part 8: ISO/IEC 23001-8 – Common Encryption for MPEG-2 Transport Stream
- Part 9: ISO/IEC 23001-9 – Carriage of Timed Metadata Metrics of Media in ISO Base Media File Format
- Part 10: ISO/IEC 23001-10 – Green Metadata
- Part 11: ISO/IEC 23001-11 – Sample Variants in the ISO Base Media File Format
- Part 12: ISO/IEC 23001-12 – White paper on Media Orchestration
- Part 13: ISO/IEC 23001-13 – Media orchestration
- Part 14: ISO/IEC 23001-14 – Partial File Format
- Part 15: ISO/IEC 23001-15 – Carriage of Web Resource in ISOBMFF

1.5.2.9 MPEG-C

MPEG-C [mpeg.chiariglione.org/standards/mpeg-c], formally, ISO/IEC 23002, is a suite of video standards that do not fall in other well-established MPEG video standards. It consists of the following parts:

- Part 1: ISO/IEC 23002-1 – Accuracy specification for implementation of integer-output IDCT
- Part 2: ISO/IEC 23002-2 – Fixed point 8x8 DCT/IDCT
- Part 3: ISO/IEC 23002-3 – Auxiliary Video Data Representation
- Part 4: ISO/IEC 23002-4 – Video Tool Library
- Part 5: ISO/IEC 23002-5 – Reconfigurable Media Coding Conformance and Reference Software
- Part 6: ISO/IEC 23002-6 – Tools for reconfigurable media coding implementations
- Part 7: ISO/IEC 23002-7 – Supplemental enhancement information messages for coded video bit-streams

1.5.2.10 MPEG-D

MPEG-D [mpeg.chiariglione.org/standards/mpeg-d], formally ISO/IEC 23003, is a suite of audio standards that do not fall in other well-established MPEG audio standards. It consists of the following parts:

- Part 1: ISO/IEC 23003-1 – MPEG Surround
- Part 2: ISO/IEC 23003-2 – Spatial Audio Object Coding
- Part 3: ISO/IEC 23003-3 – Unified Speech and Audio Coding

- Part 4: ISO/IEC 23003-4 – Audio Dynamic Range Control
- Part 5: ISO/IEC 23003-5 – Uncompressed Audio in MP4 FF

1.5.2.11 MPEG-E

MPEG-E [mpeg.chiariglione.org/standards/mpeg-e], formally ISO/IEC 23004, is a standard for an Application Programming Interface (API) of Multimedia Middleware (M3W) that can be used to provide a uniform view to an interoperable multimedia middleware platform. It consists of the following parts:

- Part 1: ISO/IEC 23004-1 – Architecture
- Part 2: ISO/IEC 23004-2 – Multimedia API
- Part 3: ISO/IEC 23004-3 – Component Model
- Part 4: ISO/IEC 23004-4 – Resource and Quality Management
- Part 5: ISO/IEC 23004-5 – Component Download Framework
- Part 6: ISO/IEC 23004-6 – Fault Management Framework
- Part 7: ISO/IEC 23004-7 – Integrity Management Framework
- Part 8: ISO/IEC 23004-8 – Reference Software and Conformance

1.5.2.12 MPEG-G

MPEG-G [mpeg.chiariglione.org/standards/mpeg-g], formally ISO/IEC 23092, is a suite of standards to provide new effective and interoperable solutions for genomic information processing applications. It consists of the following parts:

- Part 1: ISO/IEC 23092-1 – Transport and Storage of Genomic Information
- Part 2: ISO/IEC 23092-2 – Genomic Information Representation
- Part 3: ISO/IEC 23092-3 – API for Genomic Information Representation
- Part 4: ISO/IEC 23092-4 – Reference Software
- Part 5: ISO/IEC 23092-5 – Conformance
- Part 6: ISO/IEC 23092-6 – Genomic Annotation Representation

1.5.2.13 MPEG-H

MPEG-H [mpeg.chiariglione.org/standards/mpeg-h], formally ISO/IEC 23008, is a suite of standards for heterogeneous environment delivery of audio-visual information compressed with high efficiency. It consists of the following parts:

- Part 1: ISO/IEC 23008-1 – MPEG Media Transport (MMT)
- Part 2: ISO/IEC 23008-2 – High Efficiency Video Coding
- Part 3: ISO/IEC 23008-3 – 3D Audio
- Part 4: ISO/IEC 23008-4 – MMT Reference Software
- Part 5: ISO/IEC 23008-5 – HEVC Reference Software
- Part 6: ISO/IEC 23008-6 – 3D Audio Reference Software
- Part 7: ISO/IEC 23008-7 – MMT Conformance testing
- Part 8: ISO/IEC 23008-8 – HEVC Conformance testing
- Part 9: ISO/IEC 23008-9 – 3D Audio conformance testing
- Part 10: ISO/IEC 23008-10 – MPEG Media Transport FEC Codes
- Part 11: ISO/IEC 23008-11 – MPEG Composition Information
- Part 12: ISO/IEC 23008-12 – Image File Format
- Part 13: ISO/IEC 23008-13 – MMT Implementation Guidelines
- Part 14: ISO/IEC 23008-14 – Conversion and coding practices for high-dynamic-range and wide-colour-gamut video
- Part 15: ISO/IEC 23008-15 – Signaling, backward compatibility and display adaptation for HDR/WCG video

1.5.2.14 MPEG-I

MPEG-I [mpeg.chiariglione.org/standards/mpeg-i], formally ISO/IEC 23090, is a collection of standards to digitally represent immersive media. It consists of the following parts:

- Part 1: ISO/IEC 23090-1 – Immersive Media Architectures
- Part 2: ISO/IEC 23090-2 – Omnidirectional Media Format
- Part 3: ISO/IEC 23090-3 – Versatile Video Coding
- Part 4: ISO/IEC 23090-4 – Immersive Audio Coding

- Part 5: ISO/IEC 23090-5 – Point Cloud Compression
- Part 6: ISO/IEC 23090-6 – Immersive Media Metrics
- Part 7: ISO/IEC 23090-7 – Immersive Media Metadata
- Part 8: ISO/IEC 23090-8 – Network Based Media Processing
- Part 9: ISO/IEC 23090-9 – Geometry-based Point Cloud Compression
- Part 10: ISO/IEC 23090-10 – Carriage of Point Cloud Data
- Part 11: ISO/IEC 23090-11 – Implementation Guidelines for Network-based Media Processing
- Part 12: ISO/IEC 23090-12 – Immersive Video

1.5.2.15 MPEG-M

MPEG-M [mpeg.chiariglione.org/standards/mpeg-m], formally ISO/IEC 23006, is a suite of standards to enable the easy design and implementation of media-handling value chains whose devices interoperate because they are all based on the same set of technologies, especially MPEG technologies accessible from the middleware and multimedia services, also known as MPEG eXtensible Middleware (MXM). It consists of the following parts:

- Part 1: ISO/IEC 23006-1 – Architecture
- Part 2: ISO/IEC 23006-2 – MPEG Extensible Middleware API
- Part 3: ISO/IEC 23006-3 – Reference Software and Conformance
- Part 4: ISO/IEC 23006-4 – Elementary Services
- Part 5: ISO/IEC 23006-5 – Service aggregation

1.5.2.16 MPEG-U

MPEG-U [mpeg.chiariglione.org/standards/mpeg-u], formally ISO/IEC 23007, is a set of specifications of the exchange, the display, the control and the communication of widgets with other entities and for data format for advanced user interaction (AUI) interfaces to support various advanced user interaction devices. It provides a general-purpose technology with innovative functionality that enables its use in heterogeneous scenarios such as broadcast, mobile, home network and Web domains. It consists of the following parts:

- Part 1: ISO/IEC 23007-1 – Widgets
- Part 2: ISO/IEC 23007-2 – Advanced User Interaction Interface
- Part 3: ISO/IEC 23007-3 – Reference Software and Conformance

1.5.2.17 MPEG-V

MPEG-V [mpeg.chiariglione.org/standards/mpeg-v], formally ISO/IEC 23005, outlines an architecture and specifies associated information representations to enable interoperability between virtual worlds (e.g., digital content provider of a virtual world, gaming, simulation), and between real and virtual worlds (e.g., sensors, actuators, vision and rendering, robotics). It consists of the following parts:

- Part 1: ISO/IEC 23005-1 – Architecture
- Part 2: ISO/IEC 23005-2 – Control Information
- Part 3: ISO/IEC 23005-3 – Sensory Information
- Part 4: ISO/IEC 23005-4 – Virtual World Object Characteristics
- Part 5: ISO/IEC 23005-5 – Data Formats for Interaction
- Part 6: ISO/IEC 23005-6 – Common Types and Tools
- Part 7: ISO/IEC 23005-7 – Reference Software and Conformance

1.5.2.18 MPEG-DASH

MPEG-DASH [mpeg.chiariglione.org/standards/mpeg-dash], formally ISO/IEC 23009, is a suite of standards providing a solution for the efficient and easy streaming of multimedia using existing available HTTP infrastructure (particularly servers and CDNs, but also proxies, caches, etc.), i.e., dynamic adaptive streaming over HTTP. It consists of the following parts:

- Part 1: ISO/IEC 23009-1 – Media presentation description and segment formats
- Part 2: ISO/IEC 23009-2 – Reference software and conformance
- Part 3: ISO/IEC 23009-3 – Implementation guidelines
- Part 4: ISO/IEC 23009-4 – Format Independent Segment encryption and authentication
- Part 5: ISO/IEC 23009-5 – Server and network assisted DASH (SAND)
- Part 6: ISO/IEC 23009-6 – DASH with Server Push and WebSockets

- Part 7: ISO/IEC 23009-7 – Delivery of CMAF (Common Media Application Format) content with DASH
- Part 8: ISO/IEC 23009-8 – Reference Software and Conformance

1.5.2.19 MPEG-IoMT

MPEG-IoMT [mpeg.chiariglione.org/standards/mpeg-iomt], formally ISO/IEC 23093, is a set of APIs for communicating media devices (called Media Things), i.e., Internet of Media Things. It consists of the following parts:

- Part 1: ISO/IEC 23093-1 – IoMT Architecture
- Part 2: ISO/IEC 23093-2 – IoMT Discovery and Communication API
- Part 3: ISO/IEC 23093-3 – Media data formats and API
- Part 4: ISO/IEC 23093-4 – IoMT Reference Software and Conformance

1.5.2.20 MPEG-CICP

MPEG-CICP [mpeg.chiariglione.org/standards/mpeg-cicp], formally ISO/IEC 23091, is a suite of standards to specify code points for non-standard specific media formats. It consists of the following parts:

- Part 1: ISO/IEC 23091-1 – Systems
- Part 2: ISO/IEC 23091-2 – Video
- Part 3: ISO/IEC 23091-3 – Audio
- Part 4: ISO/IEC 23091-4 – Usage of video signal type code points

1.5.2.21 MPEG-MAR

MPEG-MAR [mpeg.chiariglione.org/standards/mpeg-mar], formally ISO/IEC 18039, is a Mixed and Augmented Reality Reference Model developed jointly with SC 24/WG 9. It consists of the following part:

- Part 1: ISO/IEC 18039-1 – Mixed and Augmented Reality Reference Model

1.5.3 Standards Organizations

Several organizations are responsible for developing and maintaining multimedia standards. These organizations bring together experts from various industries to develop and maintain standards that cover a wide range of multimedia aspects, including audio, video, graphics and interactive content. These bodies work collaboratively to define technical standards, file formats, compression algorithms and protocols that facilitate the creation, distribution and playback of multimedia content across diverse platforms and devices. Some of the well-known organizations, along with their activities, are discussed below.

1.5.3.1 Audio Engineering Society (AES)

The AES [aes.org] has been involved in setting technical standards for audio since 1977. The AES Standards Committee (AESSC), through a consensus system, develops and publishes a number of standards on the subject of analog and digital audio recording, transmission and/or reproduction. Notable standards include: (1) AES3 (also commonly known as AES/EBU) for digital audio interconnection (2) AES10 (also commonly known as MADI) for multichannel digital audio interconnection (3) AES11 for digital audio synchronization (4) AES31 file exchange format (5) AES53 for sending AES3 digital audio data over Asynchronous Transfer Mode networks (6) AES67 for audio over IP interoperability.

1.5.3.2 Advanced Television Systems Committee (ATSC)

The Advanced Television Systems Committee (ATSC) [aes.org] is an important organization in the field of digital television broadcasting standards, established in 1982. The ATSC is renowned for its work in shaping the digital television landscape and ensuring the seamless transition from analog to digital television. The ATSC develops and maintains a set of standards that encompass various aspects of digital television, including transmission, video coding, audio coding, and interactive television. The ATSC standards have played a critical role in enhancing the quality and efficiency of television broadcasting, leading to innovations such as high-definition television (HDTV), 4K ultra-high-definition television (UHDTV), and the advanced broadcast capabilities used for over-the-air, cable, satellite and internet-based TV services. Notable standards include: (1) A/52, Digital Audio Compression (AC-3), (2) A/53, Digital TV standard, (3) A/70, Conditional access system for terrestrial broadcast, (4) A/72, AVC video, (5) A/81, DTH satellite broadcast standard, (6) A/90, ATSC data broadcast system and (7) A/153, Mobile DTV standard.

1.5.3.3 International Commission on Illumination (CIE)

The International Commission on Illumination (CIE) [cie.co.at], established in 1913, is the global authority on the science and art of light and lighting.

This international organization, often referred to as the CIE from its French name "Commission Internationale de l'Éclairage," is dedicated to promoting the understanding and application of light and color. Headquartered at Vienna, Austria, the CIE develops and publishes standards, recommendations, and reports that encompass various aspects of light and lighting, including photometry, colorimetry, and the effects of light on human well-being. It plays a vital role in fields such as architecture, lighting design, automotive lighting and display technology. The CIE's work ensures that lighting is not only functional but also safe, energy-efficient, and conducive to human comfort and productivity. The organization's contributions extend to diverse domains, including healthcare, transportation, and environmental sustainability. Notable contributions include: (1) formalization of the CIE 1931 XYZ color space, (2) definitions of the 1931 CIE 2° standard observer, (3) definitions of the 1964 10° CIE standard observer, (4) development of the CIELAB and CIELUV color spaces.

1.5.3.4 Digital Imaging and Communications in Medicine Systems Committee (DICOM)

The DICOM Systems Committee [dicomstandard.org] is a vital component of the healthcare and medical imaging industry, responsible for developing and maintaining the DICOM standard. DICOM, short for Digital Imaging and Communications in Medicine, is a standardized protocol and file format used for the exchange and management of medical imaging data and associated information. This standard ensures that medical imaging devices, systems and software from different manufacturers can seamlessly communicate and share medical images and related data. The DICOM Standard Committee is responsible for: (1) Developing proposed standards within the scope of the committee, (2) Voting to approve proposed standards within its scope, (3) Maintaining the committee's standard by reaffirming, revising or withdrawing the standard within five years of its approval.

1.5.3.5 European Broadcasting Union (EBU)

The European Broadcasting Union (EBU) [ebu.ch] is a leading alliance of public service media organizations in Europe, founded in 1950. Headquartered in Geneva, Switzerland, the EBU represents over 100 member broadcasters across the European continent and beyond. The organization plays a crucial role in promoting and advocating for public service broadcasting, which aims to provide high-quality, diverse and impartial content for the public's benefit. The EBU serves as a platform for collaboration among its member organizations, facilitating the exchange of content, technology and best practices. Furthermore, the EBU actively engages in addressing technological and regulatory challenges in the broadcasting industry, emphasizing issues like spectrum management, digital transition and innovation in media content production and delivery. Notable contributions include: (1) The AES/EBU digital audio interface, formally known as AES3, (2) Serial and parallel interfaces for digital video (ITU-R Recommendations 601 and 656), (3) RDS – the radio data system used on FM broadcasting, (4) Digital radio (DAB) through Eureka Project 147 (5) DVB (Digital Video Broadcasting).

1.5.3.6 European Computer Manufacturers Association (ECMA)

The European Computer Manufacturers Association (ECMA) [ecma-international.org] is an international standards organization founded in 1961, headquartered in Geneva, Switzerland. ECMA plays a vital role in the development and standardization of information and communication technology (ICT) standards. Its primary mission is to foster innovation, interoperability and competitiveness in the ICT sector. ECMA is particularly known for its work in developing standards for various technologies, including computer programming languages, file formats and optical storage media. Notable contributions include: (1) ECMA -6, 7-bit coded character set, (2) ECMA-107 – FAT12/FAT16 file system, (3) ECMA-119 – CD-ROM volume and file-structure (later known as ISO 9660), (4) ECMA-130 – CD-ROM "Yellow Book" format, (5) ECMA-262 – ECMAScript Language Specification (based on JavaScript), (6) ECMA-334 – C# Language Specification, (7) ECMA-335 – Common Language Infrastructure (CLI), (8) ECMA-372 – C++/CLI Language Specification, (9) ECMA-376 – Office Open XML (10) ECMA-404 – JSON.

1.5.3.7 International Color Consortium (ICC)

The International Color Consortium (ICC) [color.org] is a global organization formed in 1993 that focuses on developing and promoting open, vendor-neutral color management standards. The ICC plays a critical role in ensuring consistent and accurate color reproduction across different devices, applications and media. Its primary goal is to create a standard framework for translating and preserving color information, which is essential in industries like graphic design, printing, photography and digital imaging. The ICC defines color profiles, such as International Color Consortium Profile (ICC profile), which provides detailed information about a device's color capabilities enabling color sensitive software to render and reproduce colors consistently. These profiles are used to maintain color fidelity from the creation of an image to its final output, be it on a computer monitor, in print or on other display devices.

1.5.3.8 International Electrotechnical Commission (IEC)

The International Electro-technical Commission (IEC) [iec.ch] is a prominent global organization dedicated to standardizing and promoting international standards in the field of electro-technology. Founded in 1906, the IEC plays an important role in advancing safety, efficiency and interoperability in electrical and electronic devices, systems, and technologies. Its broad scope encompasses a wide array of subjects, including electrical standards for appliances, power generation, renewable energy and telecommunications, among others. The IEC's comprehensive standardization work ensures that products and systems are safe to use, energy-efficient, and compatible with global markets. The IEC cooperates closely with the International Organization for Standardization (ISO) and the International Telecommunication Union (ITU). In addition, it works with several major standards development organizations, including the IEEE. Standards developed jointly with ISO, carry the acronym of both organizations. The use of the ISO/IEC prefix covers publications from ISO/IEC Joint Technical Committee 1 – Information Technology. Some notable contributions include: (1) IEC 60065 Audio, video and similar electronic apparatus – Safety requirements, (2) IEC 60118 Electro-acoustics – Hearing aids, (3) IEC 60461 Time and control code, (4) IEC 60500 Underwater acoustics – Hydrophones, (5) IEC 60558 Type C helical video tape recorders, (6) IEC 60601 Medical electrical equipment, (7) IEC 60617 Graphical symbols for diagrams, (8) IEC 60757 Color codes in electronic components, (9) IEC 60794 Optical fiber cables, (10) IEC 60908 Audio recording – Compact disc digital audio system, (11) IEC 60958 Digital audio interface, (12) IEC 61747 Liquid crystal display devices,(13) IEC 61988 Plasma display panels, (14) IEC 62272 Digital radio mondiale (DRM), (15) IEC 62680 Universal Serial Bus (USB) interfaces.

1.5.3.9 Institute of Electrical and Electronics Engineers (IEEE)

The Institute of Electrical and Electronics Engineers (IEEE) [ieee.org] is a renowned and globally influential professional organization dedicated to the advancement of technology and innovation in the fields of electrical engineering, electronics and computer science. Established in 1963 through the merger of the American Institute of Electrical Engineers (AIEE) and the Institute of Radio Engineers (IRE), IEEE has grown to become the world's largest technical professional organization. It provides a wide array of resources, services and opportunities for its members, who include engineers, scientists and technologists. IEEE's core mission involves the development and promotion of industry standards, publications, conferences and educational programs, with a particular focus on areas such as telecommunications, electronics, power and energy, and computing. The IEEE Standards Association (IEEE SA) [standards.ieee.org] is an operating unit within IEEE that develops global standards in a broad range of industries, including power and energy, artificial intelligence systems, internet of things, consumer technology and consumer electronics, biomedical and health care, learning technology, information technology and robotics, telecommunication, automotive, transportation, home automation, nanotechnology, information assurance and emerging technologies. Notable contributions include: (1) IEEE 802.3 Ethernet Standards for Carrier Sense Multiple Access with Collision Detection (CSMA/CD), (2) IEEE 802.11 Wireless Networking – "WiFi", (3) IEEE 802.8 Fiber-optic connection, (4) IEEE P1363 Public key cryptography, (5) IEEE 1394 Serial bus – "FireWire", (6) IEEE 2413 Standard for an Architectural Framework for the Internet of Things (IoT).

1.5.3.10 Internet Engineering Task Force (IETF)

The Internet Engineering Task Force (IETF) [ietf.org] is an important and highly influential open international community of network designers, operators, vendors and researchers. Founded in 1986, the IETF plays a central role in shaping the architecture, protocols and standards of the Internet. It operates under the auspices of the Internet Society (ISOC) and supervised by the Internet Society's Internet Architecture Board (IAB). It provides a collaborative and open forum for the development and evolution of internet-related technologies. The IETF's work is driven by a commitment to openness, innovation and the principle of "rough consensus." It encompasses a wide array of technical areas, including internet protocols, routing, security, privacy and internet governance. The documents that specify and discuss current and developing standards for the Internet are called Requests for Comments, or RFCs. Notable contributions include: (1) Development of protocols and architectures related to communications like HTTP, FTP, URI schemes, MIME types, authentication mechanisms, data formats, metrics, and codecs. (2) Development of Internet protocols like IP versions 4 and 6, DNS, DHCP, host and router configuration, VPNs, etc. (3) Development of network management protocols like NETCONF, SNMP, RADIUS, Diameter, and CAPWAP. (4) Development of routing and signaling protocols such as OSPF, IS-IS, BGP, RSVP-TE and LDP. (5) Development of architecture and protocols for IoT.

1.5.3.11 International Organization for Standardization (ISO)

The International Organization for Standardization (ISO) [iso.org] is a globally recognized and influential entity that develops and publishes international standards aimed at promoting quality, safety, efficiency and innovation in various industries and sectors. Founded in 1947, ISO is a nongovernmental organization comprising representatives from over 160 member countries. It collaboratively creates and maintains standards that cover a wide spectrum of domains, including technology, manufacturing, healthcare and environmental management. ISO standards offer a common language and framework for businesses, organizations, and governments worldwide, facilitating international trade, interoperability and the delivery of safe and high-quality products and services. The ISO certification process is widely used to demonstrate compliance with these standards, signaling a commitment to excellence and best practices in fields ranging from quality management (ISO 9001) to environmental management (ISO 14001). It has published over 24,500 international standards covering almost all aspects of technology and manufacturing. The organization develops and publishes standardization in all technical and nontechnical fields other than electrical and electronic engineering, which is handled by the IEC. ISO has a joint technical committee (JTC) with the International Electrotechnical Commission (IEC) to develop standards relating to information technology (IT). Known as JTC 1 and entitled "Information technology," it was created in 1987 and its mission is "to develop worldwide Information and Communication Technology (ICT) standards for business and consumer applications." Some popular standards include: (1) ISO 9000 family – Quality management (2) ISO/IEC 27000 family – Information security management (3) ISO 45001 and related standards – Occupational health and safety (4) ISO 8601 – Date and time format (5) ISO 13485 – Medical devices (6) ISO 31000 – Risk management (7) ISO 50001 – Energy management (8) ISO 9660 – ISO images for computer files in CD (9) ISO 14000 family – Environmental management (10) ISO 22000 – Food safety management (11) ISO 4217 – Currency codes (12) ISO 3166 – Country Codes.

1.5.3.12 International Telecommunication Union (ITU)

The International Telecommunication Union (ITU) [itu.int] is a specialized agency of the United Nations dedicated to the field of information and communication technologies (ICTs). Established in 1865, the ITU has played an important role in facilitating global communication and connectivity. It develops and manages international standards, regulations and policies for telecommunications and radiocommunications, ensuring that networks and technologies can interoperate seamlessly across borders. The ITU encompasses three sectors: the Radiocommunication Sector (ITU-R), which deals with spectrum management and satellite communications; the Telecommunication Standardization Sector (ITU-T), responsible for developing technical standards; and the Telecommunication Development Sector (ITU-D), which focuses on bridging the digital divide and extending access to ICTs in underserved regions. The ITU's work is fundamental in enabling the global telecommunications ecosystem, including the development of new technologies, broadband connectivity, and emerging innovations such as 5G and the Internet of Things (IoT). The ITU-T recommendations are divided into categories, namely, G: Transmission systems and media, H: Audio-visual and multimedia systems, I: Integrated services digital network (ISDN), J: transmission of multimedia signals, V: Data communication over the telephone, X: Data networks, open system communications, Notable ITU-T recommendations include: (1) G.657: single mode optical fiber cable, (2) G.711: PCM (pulse code modulation) of audio signals, (3) G.726: ADPCM speech codec, (4) G.992.5: ADSL (asynchronous digital subscriber line) broadband Internet access, (5) H.261: video compression standard, (6) H.262: video coding format or MPEG-2 part 2, (7) H.264: AVC (advanced video coding) or MPEG-4 part 10, (8) H.265: high efficiency video coding or MPEG-H part 2, (9) H.331: standards for video conferencing, (10) T.38: fax over IP network, (11) X.25: packet switched data communicated in WAN (12) X.400: message handling system (MHS).

1.5.3.13 Joint Photographic Experts Group (JPEG)

The Joint Photographic Experts Group (JPEG) [jpeg.org] is a renowned and pivotal organization in the field of digital image compression and encoding. Established in 1986, JPEG is known for its work in developing and maintaining international standards for image compression and the storage of digital images. The JPEG format, which is one of the most widely used image compression standards, has revolutionized the way images are stored and transmitted across various digital media. JPEG's standards are integral to the efficient compression and decompression of digital images, significantly reducing file sizes while preserving image quality. The organization has also developed several JPEG-related standards, including JPEG 2000 and JPEG XR, which offer enhanced compression capabilities and image quality. JPEG's work is instrumental in applications ranging from digital photography and Web graphics to medical imaging and satellite imagery, contributing to faster data transfer and the widespread sharing of images across digital platforms.

1.5.3.14 MIDI Manufacturers Association (MMA)

The MIDI Manufacturers Association (MMA) [midi.org] is a significant organization in the realm of music and audio technology, primarily known for its work in the development and promotion of the Musical Instrument Digital Interface (MIDI) standard. Established in 1984, the MMA serves as a collaborative forum for manufacturers, developers and professionals in the music industry to create and maintain MIDI standards. MIDI is a crucial protocol that allows electronic musical instruments, computers and software to communicate and synchronize with one another. It has revolutionized music production and performance, enabling musicians and composers to connect a wide range of devices, from synthesizers to digital audio workstations, and control them in a standardized way. The MMA's standards ensure that MIDI remains an adaptable and versatile tool for musicians and producers, promoting compatibility and interoperability across various hardware and software platforms. MIDI has not only had a profound impact on music creation but also extended to other applications like lighting control, video production and interactive installations.

1.5.3.15 Moving Picture Experts Group (MPEG)

The Moving Picture Experts Group (MPEG) [mpeg.chiariglione.org] and [www.mpeg.org] is a renowned and influential entity in the domain of audio and video compression and coding standards. Established in 1988, MPEG operates under the International Organization for Standardization (ISO) and the International Electrotechnical Commission (IEC). Its primary mission is to develop and maintain international standards for the compression, decompression, processing and transmission of digital audio and video content. These standards significantly impact various industries, including digital television, streaming media, multimedia applications, and digital cinema. Some of the most well-known MPEG standards include MPEG-1, responsible for the development of formats like MP3 for audio compression, and MPEG-2, widely used in DVDs and digital broadcasting. MPEG-4 and its successors have contributed to the growth of multimedia applications, including video conferencing and internet streaming. The latest standards, such as High Efficiency Video Coding (HEVC) and Versatile Video Coding (VVC), are pivotal in enabling high-quality video compression for applications like 4K and 8K video streaming. MPEG's work continues to shape the digital media landscape, enabling efficient data transmission and multimedia experiences across various platforms and devices.

1.5.3.16 National Television Systems Committee (NTSC)

The National Television Systems Committee (NTSC) [ntsc-tv.com] is a historically significant organization in the United States that played a fundamental role in the development and standardization of analog television broadcasting systems. Established in 1940, the NTSC was responsible for defining the analog broadcast television standards that shaped the American television landscape for decades. One of its most notable contributions is the creation of the NTSC color television standard, which introduced color television broadcasts in the United States. The NTSC color standard, adopted in 1953, became the foundation for color television broadcasting in North America and several other countries. This innovation revolutionized the way people experienced television, offering a richer and more immersive viewing experience. Over time, the NTSC standards became the basis for broadcasting worldwide, with minor variations and adaptations in different regions. However, as technology advanced and the transition to digital broadcasting began, the NTSC analog system was gradually phased out, and digital television standards, such as ATSC (Advanced Television Systems Committee) standards, took its place. The NTSC's work not only marked a significant milestone in the history of television but also demonstrated the importance of establishing standardized systems to ensure a consistent and high-quality viewing experience for audiences. While the NTSC's analog standards are no longer in use, its legacy remains in the form of digital television standards and the principles of standardization that continue to shape the broadcast industry today.

1.5.3.17 Open Archives Initiative (OAI)

The Open Archives Initiative (OAI) [openarchives.org] is a vital initiative in the field of digital information sharing and open access to scholarly and cultural resources. Established in 1999, OAI aims to facilitate interoperability and efficient sharing of digital content, particularly in the academic and research domains. At the core of the OAI is the development and promotion of protocols and standards that enable repositories, such as digital libraries and archives, to make their content openly accessible and discoverable on the internet. The most well-known standard developed by OAI is the OAI Protocol for Metadata Harvesting (OAI-PMH), which allows repositories to expose their metadata to others, making it easier for services and search engines to index and retrieve scholarly and cultural resources. OAI has significantly contributed to the open-access movement, ensuring that research, educational materials and cultural heritage resources are readily available to a global audience. It has also played a crucial role in fostering collaboration among institutions and organizations committed to open-access principles, ultimately promoting the dissemination of knowledge and cultural heritage in the digital age.

1.5.3.18 Radio Corporation of America (RCA)

The Radio Corporation of America (RCA) [rca.com] was one of the most influential and pioneering electronics companies in the 20th century. Established in 1919, RCA was a key player in the development of radio and television technology and a major contributor to the early days of the American broadcasting industry. The corporation's founder, David Sarnoff, played a visionary role in advocating for the widespread adoption of radio and television. RCA made several ground-breaking contributions to the world of electronics, including the development of the National Broadcasting Company (NBC), one of the major television networks in the United States. RCA also played a significant role in the development of color television, introducing the NTSC color television standard in the 1950s. RCA's inventions and innovations extended to the production of electronic equipment and components. The company was a leader in the manufacturing of radios, televisions and a wide range of electronic products. Over time, RCA evolved into a diversified conglomerate with interests in various industries, including defense electronics, aerospace and semiconductor manufacturing. The corporation's research and development efforts made substantial contributions to the advancement of electronics and telecommunications technology.

1.5.3.19 Society of Motion Picture and Television Engineers (SMPTE)

The Society of Motion Picture and Television Engineers (SMPTE) [smpte.org] is a highly respected and influential organization in the world of film, television and motion-imaging technology. Established in 1916, SMPTE has been at the forefront of the motion-imaging industry, playing an important role in standardization, innovation, and the development of technologies that have shaped the way we create, deliver, and experience motion pictures and television content. SMPTE is renowned for its work in establishing standards for the film and television industries, including frame rates, aspect ratios, color spaces, and digital cinema. Its standards have ensured consistency and compatibility in the production, distribution and display of content, and have been integral to the transition from analog to digital technology in the industry. SMPTE has also played a vital role in the development of high-definition television (HDTV) and ultra-high-definition television (UHDTV), and it continues to address emerging technologies like high dynamic range (HDR) and immersive audio. As the industry continues to evolve with innovations in areas like virtual reality, augmented reality and streaming media, SMPTE remains an essential organization in the ongoing development and standardization of technologies that enrich the world of motion pictures and television.

1.5.3.20 Unicode Consortium (UC)

The Unicode Consortium [unicode.org/consortium] is a pioneering and essential organization that has revolutionized text encoding and character standards in the digital world. Founded in 1987, the consortium is a collective effort of major technology companies, software developers, linguists and experts from around the globe. Its primary mission is to develop and maintain the Unicode Standard, a comprehensive character encoding system that has become the backbone of multilingual text representation in computing. This standard provides a consistent and universal framework for encoding and processing text in different languages, scripts and writing systems, ensuring seamless interoperability across digital platforms, devices, and applications. The Unicode Standard supports an extensive range of characters, including alphabets, ideographs, symbols, emojis and more.

It also provides guidelines for text layout, line-breaking and bidirectional text support, making it possible to handle complex scripts and diverse languages with precision.

1.5.3.21 World Wide Web Consortium (W3C)

The World Wide Web Consortium (W3C) [w3.org], founded in 1994 by internet pioneers such as Tim Berners-Lee (Berners-Lee, 1990), is a globally recognized organization dedicated to the development and standardization of Web technologies. W3C plays an important role in ensuring the interoperability and accessibility of the World Wide Web. It develops open, vendor-neutral standards and recommendations that define the building blocks of the Web, including HTML, CSS and Web accessibility guidelines. These standards are crucial in ensuring that Web content can be created, presented and accessed consistently across different platforms, browsers, and devices. Furthermore, the W3C focuses on Web security, privacy, and emerging technologies like the Web of Things (WoT) and WebAssembly. Its commitment to Web accessibility promotes an inclusive digital environment, allowing all users, regardless of abilities or disabilities, to access and engage with Web content. W3C's impact extends to Web development, design, and user experience, as well as applications like e-commerce, social media and e-learning. As the steward of Web standards, the W3C continues to foster innovation and maintain the fundamental principles of an open, accessible and universally usable World Wide Web.

1.5.3.22 Web Hypertext Application Technology Working Group (WHATWG)

The Web Hypertext Application Technology Working Group (WHATWG) [whatwg.org] is a standards body with representations from Apple, Mozilla, Google, Opera and Microsoft, responsible for maintaining multiple Web-related technical standards, including the specifications for the Hyper Text Markup Language (HTML) and the Document Object Model (DOM). The WHATWG was formed in response to the slow development of World Wide Web Consortium (W3C) Web standards and W3C's decision to abandon HTML in favor of XML-based technologies. In 2019, the W3C and WHATWG agreed to a memorandum of understanding where development of HTML and DOM specifications would be done principally in the WHATWG. The HTML specification has been designated a living document, which implies a document which is continually edited and updated.

1.6 Multimedia Technologies and Platforms

1.6.1 Frontend Technologies

Frontend technologies for multimedia production primarily involve the tools and frameworks used to create, display and interact with multimedia content on the user's side. These technologies encompass various aspects, from Web and mobile development to user interfaces and user experience design. Major frontend technologies and components for multimedia production include the following.

HTML5: HTML5 [html.spec.whatwg.org/multipage] is the standard markup language for structuring and presenting content on the Web. It includes elements and attributes for embedding multimedia content such as audio and video using the "audio" and "video" tags.

CSS3: Cascading Style Sheets (CSS) [www.w3.org/TR/CSS] is used for styling and formatting multimedia content. CSS3 introduces advanced styling features, animations and transitions that enhance the visual appeal of multimedia elements.

JavaScript: JavaScript [ecma-international.org/publications-and-standards/standards/ecma-262] is a critical programming language for multimedia interactivity and user experience enhancement. Libraries and frameworks like jQuery [jquery.com], Three.js (for 3D graphics) [threejs.org] and PixiJS (for 2D graphics) [pixijs.com] are commonly used for multimedia applications.

WebRTC: Web Real-Time Communication (WebRTC) [webrtc.org] technology enables real-time communication, including video and audio conferencing, screen sharing and peer-to-peer multimedia streaming directly in Web browsers.

Media Player Libraries: Various JavaScript libraries and frameworks provide customizable media players for Web applications. Examples include Video.js [videojs.com], Plyr [plyr.io], and JW Player [jwplayer.com].

Canvas and WebGL: The HTML5 "canvas" element and WebGL [khronos.org/webgl] are used for rendering 2D and 3D graphics, making them suitable for interactive multimedia experiences and games.

Web Audio API: The Web Audio API [w3.org/TR/webaudio] is used to create, manipulate and process audio content in real time, making it ideal for multimedia applications that involve sound generation and processing.

Responsive Design: Responsive Web design techniques and frameworks like Bootstrap [getbootstrap.com] and Foundation [get.foundation] are crucial for ensuring multimedia content displays properly on various screen sizes and devices.

UI/UX Frameworks: User interface (UI) and user experience (UX) frameworks like React [react.dev], Angular [angular.dev] and Vue.js [vuejs.org] are used for building interactive multimedia applications and providing a seamless user experience.

Mobile App Development: Technologies such as React Native [reactnative.dev] and Flutter [flutter.dev] are used for creating cross-platform mobile applications that can incorporate multimedia elements.

Augmented Reality (AR) and Virtual Reality (VR): AR and VR technologies often rely on frontend frameworks and libraries like A-Frame [aframe.io] and AR.js [ar-js-org.github.io/AR.js-Docs] for building immersive multimedia experiences.

SVG (Scalable Vector Graphics): SVG [w3.org/Graphics/SVG] is used for creating vector-based graphics and animations that can be scaled without loss of quality, making it suitable for multimedia applications.

Animation Libraries: Animation libraries like GreenSock (GSAP) [gsap.com] and Anime.js [animejs.com] are used to create dynamic and interactive animations within multimedia content. GreenSock Animation Platform (GSAP) is a robust JavaScript animation library known for its versatility and performance, including advanced features like Bezier curves, physics-based motion and precise control over timelines. Anime.js is a lightweight JavaScript animation library designed for simplicity and flexibility, and supports a variety of animation properties, easing functions and call-backs.

Font and Icon Libraries: Libraries like Google Fonts [fonts.google.com] and Font Awesome [fontawesome.com] are used to enhance the visual aspects of multimedia content by providing a wide range of fonts and icons.

UI Design and Prototyping Tools: Design and prototyping tools such as Adobe XD [adobe.com/xd], Sketch [sketch.com] and Figma [figma.com] help designers create multimedia interfaces and test user interactions. Adobe XD is a powerful design and prototyping tool tailored for user experience (UX) and user interface (UI) designers with features like responsive resizing, shared design components and collaborative design capabilities. Figma is a cloud-based design and collaboration platform that stands out for its real-time collaboration features allowing instant collaboration with team members. Sketch is a vector-based design tool tailored for macOS, widely used by UI and UX designers, known for its extensive plugin ecosystem, offering additional functionalities and integrations.

WebVR and WebXR: WebVR [webvr.info] and WebXR [w3.org/TR/webxr] technologies allow the creation of VR and AR experiences directly within Web browsers, enabling immersive multimedia content.

These frontend technologies are essential for creating engaging and interactive multimedia content for the Web and mobile platforms. They provide the tools and frameworks necessary to design, develop and deliver multimedia applications and experiences to users.

1.6.2 Backend Technologies

In multimedia production, backend technologies play a crucial role in managing, processing and delivering multimedia content efficiently. Some major backend technologies and components commonly used in multimedia production are discussed below.

Media Servers: Media servers are essential for streaming audio and video content. They handle tasks such as content delivery, load balancing, and real-time transcoding. Popular media server software includes Wowza Streaming Engine [wowza.com/streaming-engine], Adobe Media Server [business.adobe.com/products/primetime/adobe-media-server.html] and Red5 [red5server.com].

Content Management Systems (CMS): CMS platforms are used to organize, store, and manage multimedia assets. They provide tools for content creation, version control, metadata management and asset organization. Examples include WordPress [wordpress.org], Adobe Experience Manager [business.adobe.com/products/experience-manager/adobe-experience-manager] and Drupal [drupal.org].

Media Encoding and Transcoding: Media encoding and transcoding tools are used to convert multimedia content from one format to another, ensuring compatibility and optimizing file size for various devices and platforms. FFmpeg [ffmpeg.org] and HandBrake [handbrake.fr] are widely used tools for this purpose.

Database Management Systems (DBMS): DBMS is used to store metadata, organize content libraries, and manage user access rights. Popular DBMS options include MySQL [mysql.com], PostgreSQL [postgresql.org], MongoDB [mongodb.com] and Microsoft SQL Server [microsoft.com/en-us/sql-server].

Digital Asset Management (DAM) Systems: DAM systems focus on organizing and distributing multimedia assets. They provide features for asset metadata, version control, access control and asset distribution. Examples include Adobe Experience Manager [business.adobe.com/products/experience-manager/adobe-experience-manager], Widen Collective [acquia.com/products/acquia-dam] and Bynder [bynder.com].

Content Delivery Networks (CDNs): CDNs are used for efficient content distribution by replicating multimedia content across geographically distributed servers. CDN providers include Akamai [akamai.com], Cloudflare [cloudflare.com] and Amazon CloudFront [aws.amazon.com/cloudfront].

Video and Audio Editing Tools: Backend servers may integrate video and audio editing tools for post-production tasks, including editing, transcoding and rendering. Adobe Premiere Pro [adobe.com/products/premiere], Final Cut Pro [apple.com/final-cut-pro] and Avid Media Composer [avid.com/media-composer] are examples of video editing software.

Machine Learning and AI: Backend technologies leverage machine learning and artificial intelligence for content analysis, automatic tagging, facial recognition and content recommendation. TensorFlow [tensorflow.org], PyTorch [ai.meta.com/tools/pytorch] and OpenCV [opencv.ai] are commonly used libraries for AI and ML in multimedia production.

Cloud Computing Services: Cloud platforms such as Amazon Web Services (AWS) [aws.amazon.com], Microsoft Azure [azure.microsoft.com] and Google Cloud [cloud.google.com] offer a wide range of tools and services for multimedia processing, storage, and delivery.

Security and Rights Management: Solutions for digital rights management (DRM) like Fasoo [en.fasoo.com/strategies/enterprise-drm] and encryption are used to protect multimedia content from piracy and unauthorized access. Fasoo provides data security solutions, with advanced tools to protect, manage and control sensitive digital assets, using a file-centric security model, ensuring that security measures are embedded directly within the files themselves.

Content Analytics: Backend technologies for multimedia production often include analytics tools that track user engagement, content performance and audience behavior. Google Analytics [analytics.google.com], Adobe Analytics [business.adobe.com/products/analytics] and custom analytics solutions are used for this purpose. Google Analytics enables the tracking of a wide array of metrics related to website and app performance, from visitor traffic and user engagement to conversion rates and e-commerce transactions, the platform provides a comprehensive overview of how users interact with online content. Adobe Analytics excels in providing a unified view of customer interactions across multiple channels, including websites, mobile apps, social media and more; by aggregating data from various touch-points, businesses gain a comprehensive understanding of the customer journey.

These backend technologies, work in tandem to facilitate the creation, management, and distribution of multimedia content, ensuring it reaches the intended audience while maintaining quality, security and efficiency.

1.6.3 Multimedia Platforms

Multimedia platforms are integral to the creation, distribution and consumption of multimedia content. These platforms encompass a wide range of technologies, applications and services that enable the seamless integration of multimedia elements, such as images, audio, video, animations and interactive experiences, into various digital environments. From social media networks and streaming services to interactive Web applications and virtual reality experiences, multimedia platforms are the driving force behind engaging and immersive content delivery.

Social Media Networks: Platforms like Facebook [facebook.com], Instagram [instagram.com], Twitter/X [twitter.com] and TikTok [tiktok.com] enable users to create and share multimedia content, fostering social interactions through images, videos, live streams and stories.

Streaming Services: Platforms like YouTube [youtube.com], Netflix [netflix.com], Spotify [open.spotify.com] and Twitch [twitch.tv] provide on-demand access to multimedia content, including videos, music, podcasts and live broadcasts, enhancing user entertainment options.

E-commerce Platforms: Platforms like Amazon [amazon.com], Shopify [shopify.com] and WooCommerce [woo.com] incorporate multimedia features to showcase products through images, videos and 360-degree views, enhancing the shopping experience. They leverage high-quality product imagery and videos to provide customers with a detailed and immersive view of the products, typically a gallery of images from various angles, zoom capabilities and sometimes product demonstration videos.

Gaming Consoles and Engines: Platforms like PlayStation [www.playstation.com], Xbox [www.xbox.com], Steam [store.steampowered.com], Unity [unity.com] and Unreal Engine [www.unrealengine.com] offer multimedia capabilities for gaming, including immersive 3D graphics, spatial audio and streaming services.

Augmented Reality (AR) and Virtual Reality (VR): AR and VR platforms such as ARKit [developer.apple.com/augmented-reality/arkit], Oculus [www.oculus.com] and HTC Vive [www.vive.com] create immersive multimedia experiences with 3D audio, video and interactive elements.

1.6.4 Transport Mechanisms

Multimedia content delivery over networks is a crucial aspect of the multimedia landscape, enabling the seamless distribution of multimedia content to users around the world.

1.6.4.1 Communication Protocols

Hypertext Transfer Protocol (HTTP): HTTP [httpwg.org/specs] is the foundation of the World Wide Web and is widely used for delivering multimedia content over the internet. HTTP is essential for accessing Web pages and multimedia files, making it one of the most common protocols for content delivery.

Hypertext Transfer Protocol Secure (HTTPS): HTTPS is the secure version of HTTP, using encryption to protect the integrity and confidentiality of multimedia content during transmission. It is especially crucial for e-commerce, online streaming and secure communication.

File Transfer Protocol (FTP): FTP [datatracker.ietf.org/doc/html/rfc114] is an older protocol used for transferring multimedia files between a client and a server. While less commonly used for Web content delivery, FTP remains relevant for managing large multimedia files and server-to-server transfers.

Real-Time Transport Protocol (RTP): RTP [datatracker.ietf.org/doc/html/rfc3550] is specifically designed for real-time multimedia data transmission, making it vital for video and audio streaming, video conferencing and online gaming. RTP is often used in combination with Real-Time Control Protocol (RTCP) [datatracker.ietf.org/doc/html/rfc5760] to monitor data delivery and quality.

Internet Protocol (IP) and Transmission Control Protocol (TCP): IP [datatracker.ietf.org/doc/html/rfc791] and TCP [datatracker.ietf.org/doc/html/rfc9293] are fundamental to internet communication. IP handles routing and addressing, while TCP ensures reliable data delivery. These protocols are essential for multimedia content delivery, ensuring that data packets reach their destination accurately.

1.6.4.2 Streaming Protocols

HTTP-based streaming, using protocols like HTTP Live Streaming (HLS) [developer.apple.com/streaming/] and Dynamic Adaptive Streaming over HTTP (DASH) [www.iso.org/standard/83314.html], is a widely used method for delivering multimedia content. It allows adaptive streaming, optimizing quality based on the viewer's internet connection.

Real-Time Streaming Protocol (RTSP): RTSP [https://datatracker.ietf.org/doc/html/rfc2326] is designed for real-time streaming and controlling the delivery of multimedia content. It is commonly used for video surveillance systems, video conferencing and real-time applications.

Peer-to-Peer (P2P) Streaming: P2P streaming leverages the resources of viewers, distributing multimedia content efficiently. This technology is used in applications like live streaming, where multiple users share the streaming load.

Progressive Download: Progressive download allows multimedia content to start playing as soon as a portion of the file is downloaded. It is often used for Web video players, making content accessible even before the entire file is received.

WebRTC (Web Real-Time Communication): WebRTC [https://www.w3.org/TR/webrtc/] enables real-time communication directly between Web browsers. It is commonly used for video conferencing, live streaming and interactive multimedia applications.

Content Delivery Networks (CDNs): CDNs are networks of servers distributed globally to deliver multimedia content efficiently. They optimize content delivery by caching data closer to end-users, reducing latency and bandwidth usage.

VoIP (Voice over Internet Protocol): VoIP technology supports real-time voice communication over the internet. It is used in multimedia applications like online calls, video conferencing and voice chat in online gaming.

Real-Time Messaging Protocols: Real-time messaging protocols, such as MQTT [mqtt.org] and XMPP [xmpp.org], enable real-time communication for messaging and presence information. These protocols are essential for chat applications and multimedia messaging.

1.7 Multimedia and Games

Multimedia and games are deeply intertwined, as multimedia elements play a fundamental role in creating engaging, interactive and immersive gaming experiences. The integration of multimedia elements, such as graphics, audio, video and animations, is essential for enhancing the visual and auditory aspects of games and immersing players in virtual worlds (Gazaway, 2021).

1.7.1 Role of Multimedia in Games

Graphics: High-quality graphics are a cornerstone of modern games. Graphics encompass 2D and 3D visuals, character models, environments, textures and special effects. Multimedia elements, such as images and videos, are used to create captivating in-game visuals that range from realistic to stylized art.

Audio: Audio is crucial for setting the mood and creating an immersive gaming experience. Sound effects, background music, voice acting and spatial audio contribute to a game's atmosphere and storytelling.

Video: Video sequences, including cut-scenes and cinematics, help convey the game's narrative and add cinematic flair. Videos can be pre-rendered or displayed in real time within the game engine.

Animation: Animations breathe life into characters and objects within the game. Multimedia elements, such as animated sprites and character models, enable fluid movements and realistic interactions.

UI/UX Design: User interfaces (UI) in games benefit from multimedia elements, incorporating images, icons, animations and video to create intuitive and engaging menus and HUDs (Heads-Up Displays).

Storytelling: Multimedia elements, such as voice-overs and visual narratives, facilitate storytelling and character development within games.

1.7.2 Impact on Player Engagement

Immersive Experience: High-quality graphics and audio contribute to a more immersive gaming experience, drawing players into the game world.

Emotional Connection: Multimedia elements can evoke emotions, making players more invested in the game's characters and storyline.

Interactivity: Games utilize multimedia to create interactive environments where players can explore, solve puzzles and make choices.

Atmosphere: Soundtracks and audio effects set the tone for different game situations, from tense battles to tranquil exploration.

Feedback: Visual and auditory cues provide feedback to players, helping them understand the game's mechanics and progress.

1.7.3 Technologies

Game Engines: Game engines like Unity [unity.com], Unreal Engine [unrealengine.com] and Godot [godotengine.org] provide a framework for creating games with integrated multimedia elements. They offer tools for graphics, audio, animations and scripting. Unity is a powerful and versatile game development engine that has evolved into a comprehensive platform for creating a wide range of digital experiences, from video games to augmented reality (AR), virtual reality (VR) and simulations with extensive support for cross-platform compatibility. Unreal Engine is a leading game development framework renowned for its cutting-edge graphics, robust features with advanced rendering capabilities, including real-time global illumination, dynamic lighting and sophisticated texture mapping. Godot Engine, an open-source game development framework, has gained prominence for its versatility, ease of use and collaborative community-driven approach which includes a visual scripting system that allows developers to create game logic through a node-based interface without writing traditional code.

3D Modeling and Animation Software: Tools like Blender [www.blender.org], Maya [www.autodesk.in/products/maya] and 3ds Max [www.autodesk.com/products/3ds-max] are used to create 3D models, characters and animations.

Sound Design Software: Audio professionals use software like Adobe Audition [www.adobe.com/products/audition], Pro Tools [www.avid.com/pro-tools] and Audacity [www.audacityteam.org] to create and edit sound effects and music for games.

Video Editing Software: Video sequences and cut-scenes are often created and edited using software like Adobe Premiere Pro [www.adobe.com/products/premiere], Final Cut Pro [www.apple.com/final-cut-pro] and Sony Vegas [www.vegascreativesoftware.com/us/vegas-pro].

Texture and Image Editing Software: Tools like Adobe Photoshop [www.adobe.com/products/photoshop] and GIMP [www.gimp.org] are used to create and edit textures and visual assets for games.

Interactive Storytelling Tools: Game development platforms like Twine [twinery.org] and Ren'Py [www.renpy.org] facilitate the creation of interactive narratives in games. Twine is an open-source tool that has gained popularity for its simplicity and effectiveness in creating interactive stories and text-based games, through decision trees, enabling writers to create multiple paths and outcomes based on user choices. Ren'Py is a free and open-source visual novel engine designed to facilitate the creation of interactive storytelling experiences using a scripting language to define characters, dialogue, choices and branching narratives.

1.7.4 Mobile Gaming

Mobile phone games have become an integral part of the modern digital entertainment offering diverse genres and experiences to players of all ages. One of the key strengths of mobile phone games lies in their accessibility – since nearly everyone owns a smartphone, gaming is no longer restricted to dedicated gaming consoles or PCs. Players can enjoy games anytime, anywhere, making mobile gaming a convenient and versatile form of entertainment. An extensive variety of game genres and gameplay styles ensure that players can find experiences that align with their interests and preferences. Many mobile games operate on a freemium model, allowing players to download and play for free while offering in-app purchases for additional content, virtual items or premium features. This model has proven successful, enabling developers to reach a broad audience and monetize their games through various strategies. In addition to in-app purchases, advertising, sponsorships and subscription services contribute to the monetization scenario of mobile games. Moreover, mobile games often incorporate social and multiplayer features, fostering a sense of community and interaction among players. Integration with social media platforms further allows players to share achievements, scores, and experiences, creating a connected gaming community. In recent times augmented reality (AR) and virtual reality (VR) technologies are also making their way into mobile gaming, promising new dimensions of immersive gameplay.

1.8 Multimedia and Interactivity

Multimedia and interactivity are intertwined in various digital experiences, from websites and applications to e-learning modules and entertainment content. Interactivity refers to the ability of users to engage with digital content, while multimedia encompasses visual, audio and interactive elements (Isaias & Blashki, 2014).

1.8.1 User Engagement

Visual Engagement: Multimedia elements, such as images, videos, animations and graphics, capture users' attention and convey information more effectively than text alone.

Audio Enhancement: Sound and music enrich the user experience, setting the mood, conveying emotions and providing auditory feedback.

Interaction: Interactivity allows users to participate actively rather than passively consume content. It can take the form of clickable buttons, interactive widgets, drag-and-drop actions and more.

Feedback Mechanism: Interactivity enables real-time feedback. Users receive responses to their actions, such as confirmation messages, error alerts or dynamic content updates.

Personalization: Interactive multimedia can adapt to users' preferences and behaviors, creating a more personalized and engaging experience.

1.8.2 Role of Interactivity

Enhanced Learning: In e-learning, interactive multimedia engages learners, making the learning process more effective and enjoyable.

Storytelling: Multimedia and interactivity bring narratives to life, making storytelling in video games, virtual reality experiences and interactive websites more immersive.

Engaging Presentations: Interactive multimedia presentations are more compelling and memorable, making them effective in business, education and entertainment.

Gamification: Gamification elements in multimedia content enhance user engagement by introducing game-like features, such as points, rewards and challenges.

1.8.3 Technologies

JavaScript: JavaScript is a versatile language for creating interactive multimedia elements on websites, such as image sliders, video players and interactive forms.

HTML5: HTML5 offers native support for multimedia elements like video and audio, making it easier to integrate and control interactive multimedia.

CSS Animations: CSS animations and transitions bring interactivity to Web design, allowing for hover effects, smooth transitions and interactive buttons.

WebGL: WebGL enables 3D graphics and interactivity in Web applications, leading to more immersive experiences.

e-Learning Authoring Tools: Tools like Articulate Storyline, Adobe Captivate and Moodle [moodle.org] enable the creation of interactive e-learning modules with multimedia components.

Game Engines: Game engines like Unity [unity.com] and Unreal Engine [www.unrealengine.com] provide the foundation for creating highly interactive multimedia-rich video games.

1.8.4 Interactive Elements

Buttons and Hyperlinks: Buttons and hyperlinks are fundamental interactive elements in multimedia. Users can click or tap on them to navigate between different sections of the content or access additional information. They are commonly used in websites, e-books and presentations.

Forms and Input Fields: Forms with input fields allow users to submit data, such as text, numbers or selections, into the multimedia application. These are frequently used for user registrations, surveys and e-commerce transactions.

Sliders and Scrollers: Sliders and scrollers enable users to interact with multimedia content by adjusting values or scrolling through information. They are used in settings like audio and video players, interactive info-graphics and product galleries.

Interactive Images: Images can become interactive by incorporating hotspots or clickable areas that reveal additional information or trigger actions when selected. Interactive images are valuable for educational content, product catalogs and storytelling.

Animations and Transitions: Multimedia elements like animations, transitions and animated characters respond to user input or time-based triggers. These enhance storytelling, user engagement and interactive learning experiences.

Games and Simulations: Interactive multimedia elements are prominently featured in gaming and simulation applications. Users can control characters, make decisions, and influence outcomes, resulting in immersive gaming and training experiences.

Interactive Video: Interactive video allows users to make choices within the video, leading to different storylines or outcomes. This is commonly used in interactive movies, training videos and marketing campaigns.

1.8.5 Application Areas

E-Learning and Education: Interactive multimedia elements play a vital role in e-learning platforms, where they facilitate quizzes, simulations and personalized learning experiences. They cater to diverse learning styles and engage students effectively.

Entertainment and Gaming: Games and entertainment applications heavily rely on interactive multimedia elements to create immersive experiences. Gamers can control characters, solve puzzles and explore virtual worlds, fostering engagement and enjoyment.

Marketing and Advertising: Interactive multimedia elements are used in marketing campaigns and advertisements to encourage user participation. Interactive banners and product configurators enable users to interact with brands and products.

Museums and Exhibitions: In cultural institutions and exhibitions, interactive multimedia elements enhance visitor experiences. Touchscreen kiosks, augmented reality displays, and interactive installations provide in-depth information and engagement.

Healthcare and Simulations: Healthcare training and simulations use interactive multimedia elements to simulate medical procedures, patient scenarios and surgical training. These elements ensure realistic and hands-on learning experiences.

1.8.6 Challenges and Considerations

Usability and Accessibility: Designing interactive multimedia elements that are user-friendly and accessible to a broad audience is a challenge. Ensuring that elements are intuitive and comply with accessibility standards is essential.

Performance Optimization: Interactive multimedia elements can be resource-intensive. Optimizing performance to ensure smooth user experiences, particularly on mobile devices and Web applications, requires careful consideration.

Content Integration: Seamlessly integrating various interactive elements within multimedia applications while maintaining a coherent user experience can be complex. Synchronization, data formats and content presentation are critical considerations.

Cross-Platform Compatibility: Ensuring that interactive multimedia elements function consistently across different devices and platforms requires extensive testing and adaptation.

1.9 Multimedia Content Creation

Multimedia content creation typically follows a structured workflow. While specific workflows may vary depending on the project and organization, a general outline includes the following stages of content design, implementation and delivery (Assuncao & Gotchev, 2019).

1.9.1 Workflow

Concept and Planning: The process begins with the conceptualization of the multimedia project. This stage involves defining objectives, target audiences, and the overall content strategy. Creators establish a clear vision and outline the scope of the project, and also specify the objectives of the project, usually in terms of learning outcomes.

Story, Scriptwriting and Storyboarding: For multimedia projects a story outlines the entire content to be represented within the presentation. The story is designed keeping in mind the target audience and the objectives expected to be fulfilled. A well-crafted script outlines the narrative, dialogue and structure. It specifies which media element is used to represent which portions of the story, and how text image audio video should be synchronized with each other to deliver an integrated presentation. Storyboarding is a visual representation of the script, helping to plan the visual elements, layout, placements and orientations. It breaks up the content into individual screens and specifies how each screen should look like.

Content Creation: This stage focuses on the actual creation of multimedia assets. Depending on the project, this can involve writing text, capturing images, recording audio, shooting video or creating animations. Content creators use various tools and techniques specific to their medium.

Editing and Post-Production: Editing is a fundamental step for multimedia content refinement. For video and audio, post-production involves tasks like trimming, adding transitions, color correction and sound mixing. In image and animation projects, post-production entails image manipulation and compositing.

Integration: Multimedia content integration is the process of combining text, images, audio, video and animations into a cohesive final product. It involves careful synchronization and alignment of multimedia elements to achieve the desired user experience.

Quality Assurance and Testing: Before publishing or sharing multimedia content, it undergoes quality assurance and testing. This phase includes checking for errors, broken links and ensuring that the content performs as intended across various devices and platforms.

Distribution and Delivery: The final stage involves distributing multimedia content to the intended audience. This may involve uploading content to websites, streaming platforms, or packaging content for various mediums like CDs, DVDs or digital downloads.

1.9.2 Implementation Techniques

The techniques employed for creating multimedia content vary based on the medium. Here are some key techniques for text, image, audio, video and animation:

1.9.2.1 Text

Copywriting: Effective copywriting techniques are crucial for crafting engaging text content. This involves using persuasive language, storytelling and maintaining clarity and conciseness.

Typography: Proper selection of fonts, font sizes, line spacing and text formatting is vital for text readability and aesthetics.

1.9.2.2 Image

Photography: Techniques for capturing high-quality images include understanding lighting, composition and camera settings.

Editing: Techniques for image editing include tonal correction, color correction, sharpening, resizing and cropping, removing noise and applying filters using software like Adobe Photoshop.

1.9.2.3 Graphics

Illustration: Digital illustration techniques using software like Adobe Illustrator and Adobe Animate enable the creation of custom graphics and vector images.

1.9.2.4 Audio

Recording: Audio recording techniques encompass microphone selection, positioning, and sound treatment to capture clean and clear audio.

Editing: Techniques for audio editing include noise reduction, equalization and audio effects processing using software like Adobe Audition or Audacity.

1.9.2.5 Video

Filmmaking: Understanding camera operation, shot composition, lighting and framing are key to creating compelling video content.

Editing: Techniques for video editing involve timeline management, transitions, color grading and special effects using software like Adobe Premiere Pro or Final Cut Pro.

1.9.2.6 Animation

Keyframing: Animation techniques often involve keyframing, where key poses or states are defined, and the software interpolates between them to create smooth animations.

Rigging: In 3D animation, rigging is the process of creating skeletons and controls for characters, enabling realistic movement.

1.10 Theorems and Mathematical Formulations

1.10.1 Nyquist's Sampling Theorem

An important theorem frequently used in digital signal processing is the Nyquist Sampling Theorem named after Swedish-American engineer and physicist Harry Nyquist (1889–1976) who proposed it while investigating the bandwidth requirements for transmitting information through a telegraph channel (Nyquist, 1928). Sampling is a step for converting a continuous analog signal to a discrete set of values by examining the magnitude of the signal at specific points in time or space. The number of times sampling is done per second is called the sampling frequency. The theorem states that "When converting an analog signal into digital form, the sampling frequency used must be at least twice the bandwidth of the input signal in order to be able to reconstruct the original signal accurately from the sampled version." For a baseband signal, the bandwidth is the same as the highest frequency component, and so the theorem is often stated as the sampling frequency needs to be twice the highest frequency on the input signal. This condition is often referred to as the Nyquist criterion. If the sampling frequency falls below this threshold, then the reconstruction of the signal exhibits imperfections or artifacts, which are referred to as aliasing defects.

Thus, if f be the frequency of the analog wave and F be the sampling frequency required to digitize it, we need to have

$$F \geq 2f$$

The following figure illustrates the Nyquist's theorem by progressively sampling a sinusoidal waveform at low frequencies to generate a staircase waveform. The first figure in the top row shows the original waveform, and the second row depicts the waveform after sampling more than ten times per cycle. The third row depicts the limiting condition when the original waveform is sampled at exactly twice the frequency, i.e., two times per cycle at $\pi/2, 3\pi/2, 5\pi/2, 7\pi/2, 9\pi/2$, and $11\pi/2$ over a period of three cycles or 6π. This plot shows that since the sampling is done once in the positive half and once in the negative half of each cycle, the generated staircase waveform has at least some information to represent the two halves, i.e., their approximate durations and approximate values. The fourth row depicts the case when sampling rate is equal to the frequency of the wave; i.e., sampling is done once in each cycle. The resulting wave is a low-frequency wave which contains no information of the original wave cycles. This distortion is referred to as aliasing and occurs when the sampling rate falls below twice the wave frequency, which illustrates the Nyquist theorem (Figure 1.1)

The sampling rate and bit-depth together decide the quality of the output signal. As the sampling rate is increased, more information about the analog wave is obtained, which correspondingly increases its quality. However, more information also implies more storage space is required and powerful processors are needed to deal with them, which increases the cost. To determine a set of digitization parameters like sampling rate and bit-depth, one is required to strike a balance between the quality required and the cost which can be afforded. If a high sample rate and low bit-depth is used, then it will not be possible to represent all the sample values, hence the advantage of using a higher sampling rate will be lost. On the other hand, if a high bit-depth and low sample rate is used then there will be provisions for representing a large number of samples but actually only a fraction of it will be used, because the low sampling rate will not typically generate enough samples to fill up all the empty slots, hence, the advantage of using a higher bit-depth will be lost. Thus, in conclusion, it could be stated that sampling rate and bit-depth are independent of each other because each could be adjusted without affecting the other; however, both of them play a combined role in deciding the final quality of the digital output.

During the digitization process, a lot of approximations are used at various stages. Initially, during sampling, values of the wave only at discrete points of time are considered, again during the quantization stage, sample values are rounded off to limited number of levels, hence, errors are introduced between the digital output and analog input. Specifically, the error introduced at the quantization stage is referred to as the quantization error. The maximum error is measured as being equal to one step size of the digital output and this distortion effect is referred to as noise, since the error will randomly vary from sample to sample (in electrical circuits any signal whose amplitude varies randomly with time is referred to as "noise"). To eliminate noise fully during digitization, it must sampled at an infinite rate which is impossible in practice, hence, because of the quantization error, digitization always leads to a degradation of the original analog signal. This degradation is usually compensated at a later stage through the use of editing software.

Signal-to-Noise Ratio (SNR) is a measure of the quality of a signal and is closely related to quantization

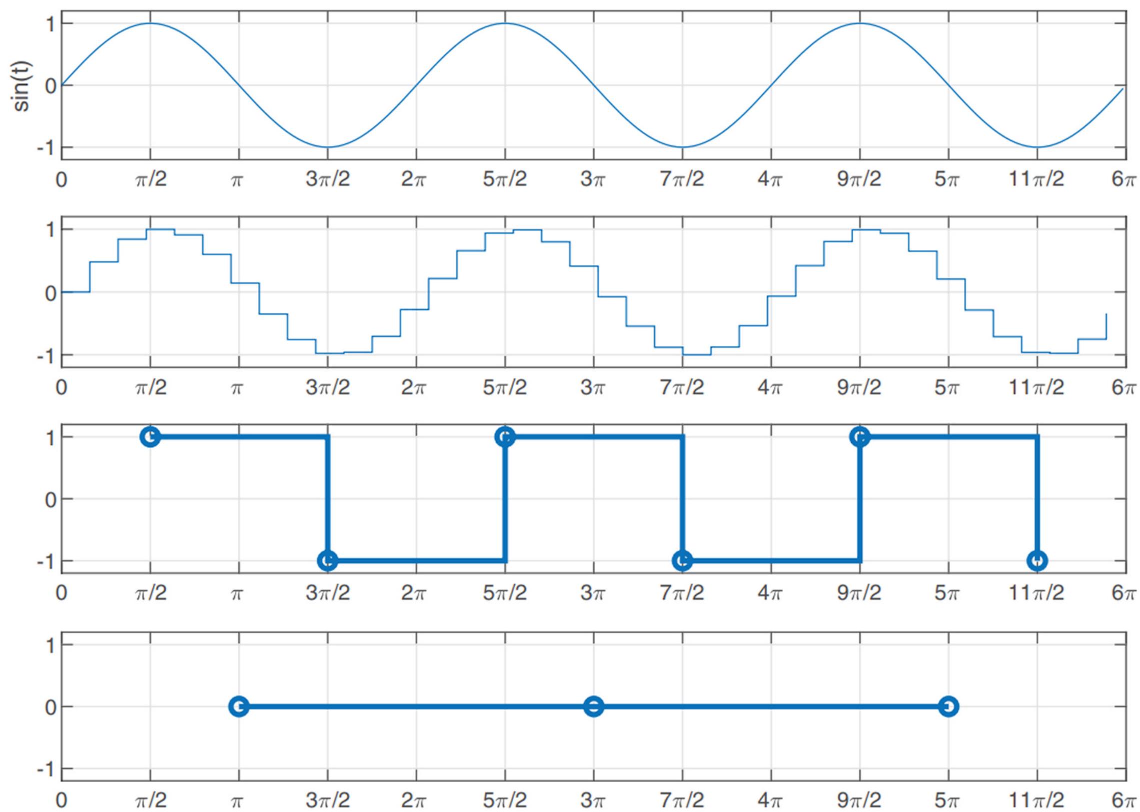

FIGURE 1.1
Illustration of Nyquist theorem.

error. SNR quantifies the ratio of amplitude of the signal of interest (A_{signal}) to the amplitude of the noise (A_{noise}) during the digitization process.

$$\text{SNR} = \frac{A_{\text{signal}}}{A_{\text{noise}}}$$

If n be the bit-depth then number of steps is 2^n. If S be peak-to-peak height of the wave, then height of each step is $S/2^n$ which is a measure of the quantization error e.

$$\text{SNR} = \frac{S}{e} = 2^n$$

Because for many signals, there is a wide difference between these two power values, SNR is often expressed in a logarithmic scale and the unit of SNR is called decibel (dB).

$$\text{SNR}_{dB} = 20\log_{10}\left(A_{\text{signal}} / A_{\text{noise}}\right)$$

$$\text{SNR}_{dB} = 20\log_{10}\left(2^n\right) = 20n(0.3) = 6n$$

The above implies that that if bit-depth is increased by 1 during digitization, the signal-to-noise ratio increases by 6 dB.

1.10.2 Euler's Theorem

Leonhard Euler (1707–1783) was a Swiss mathematician, physicist and engineer who made significant contributions to various fields of mathematics. Euler's theorem can be expressed as:

$$e^{i\pi} + 1 = 0$$

where e is the base of the natural logarithm (approximately 2.71828), i is the imaginary unit $\sqrt{-1}$, π represents the mathematical constant pi (approximately 3.14159) (Euler, 1748),

Rewriting we get:

$$e^{i\pi} = -1 = -1 + i(0) = \cos\pi + i.\sin\pi$$

Generalizing:

$$e^{ix} = \cos x + i.\sin x$$

Multimedia

Adornato, A. , Mobile and social media journalism: a practical guide for multimedia journalism, Routledge, 2021.
Assuncao, P. A. , A. Gotchev , 3D visual content creation, coding and delivery, Springer Nature Switzerland AG, 2019
Bansal, J. C. , M. S. Uddin , Computer vision and machine learning in agriculture, Springer, 2022.
Berners-Lee, T. & M. Fischetti , Weaving the Web : The original design and ultimate destiny of the World Wide Web by its inventor, Harper Collins, 1999 https://www.w3.org/People/Berners-Lee/Weaving/Overview.html
Chen, J. Y. C. , G. Fragomeni (eds), 12th International conference on virtual, augmented and mixed reality design and interaction, Springer, 2020.
Chen, S.-C. , M.-L. Shyu , Multimedia data engineering: applications and processing, Idea Group US, 2013.
Clark, R. C. , R. E. Mayer (eds), E-Learning and the science of instruction: proven guidelines for consumers and designers of multimedia learning, Wiley, 2016.
Euler, L. Introductio in analysin infinitorum, Marcum-Michaelem Bousquet, vol. 1, pp. 1–320, 1748.
Fourier, J. B. J. , Theorie analytique de la chaleur, Chez Firmin Didot, 1822.
Gadsden, D. , M. Artz , Preserving our planet: GIS for conservation, ESRI Press, 2023.
Gazaway, D. , Introduction to game systems design Pearson Education (US), 2021
Hernandez, R. K. , J. Rue , Principles of multimedia journalism: Packaging digital news, Routledge, 2015.
Hua, X.-S. , Online multimedia advertising techniques and technologies, IGI Global, 2010.
Isaias, P. , Blashki, K. , Human computer interfaces and interactivity, Idea Group US, 2014.
Jong, T. , L. Sarti (ed), Design and production of multimedia and simulation based learning material, Springer, 2012.
Mallat, S. G. , A theory for multi-resolution signal decomposition: the wavelet representation, IEEE Pattern Recognition and Machine Intelligence, vol. 11, no. 7, pp. 674–693, 1989.
Marquis, O. , B. Furht , Content-based image and video retrieval, Springer, 2012.
Mehta, M. , C. Gaie (eds), Recent advances in data and algorithms for e-Government, Springer, 2023.
Mhiripiri, N. A. , T. Chari , Media law, ethics and policy in the digital age, IGI Global, 2017.
Milutinovic, V. , Infrastructure for electronic business on the Internet, Springer, 2013.
Moreno, R. , Optimising learning from animations by minimising cognitive load: Cognitive and affective consequences of signalling and segmentation methods", Applied Cognitive Psychology, vol. 21, no. 6, pp. 765–781, 2007.
Nunes, N. , I. Oakley & V. Nisi (eds), 10th international conference on interactive digital storytelling (ICIDS), Springer, 2017.
Nyquist, H. , Certain topics in telegraph transmission theory, Transactions of American Institute of Electrical Engineers (AIEE), vol. 47, no. 2, pp. 617–644, 1928.
Pirolli, B. , Travel journalism: informing tourists in the digital age, Routledge, 2018.
Sharda, R. , D. Delen & E. Turban , Business intelligence and analytics: systems for decision support, Pearson, 2018.
Solanki, A. , S. Kumar & A. Nayyar , Handbook of research on emerging trends and applications of machine learning, IGI Global, 2019.
Steinkuehler, C. , K. Squire & S. Barab , Games, learning, and society: learning and meaning in the digital age, Cambridge University Press, 2012.
Steinmetz, R. & K. Nahrstedt , 'Multimedia computing: Prologue and overview, Communications of the ACM, vol. 38, no. 11, pp. 20–30, 1995.
Steinmetz, R. & K. Nahrstedt , Multimedia fundamentals: Media coding and content processing. Prentice Hall, 2004.
Sutherland, Ivan E. , *Sketchpad : A man-machine graphical communication system*, Technical Report UCAM-CL-TR-574, University of Cambridge, 1963 https://www.cl.cam.ac.uk/techreports/UCAM-CL-TR-574.html
Tatipamula, M. , B. Khashnabish , Multimedia communications networks: technologies and services, Artech Publishers, 1998.
Tyagi, A. K. , A. Abraham & A. Kaklauskas (eds), Intelligent interactive multimedia systems for e-healthcare applications, Springer, 2021
Valdez, G. Multimedia in education: designing complex systems for learning, IGI Global, 2006.
Vatsky, S. , Interactive museum tours: a guide to in person and virtual experiences, Rowman and Littlefield, 2023
Vaughan, T. Multimedia: Making it work, McGraw-Hill Education, 2014

Text

Boss, S. , J. C. Teague , The new Web typography, CRC Press, 2017
Chaudhuri, B. B. , Digital document processing: Major directions and recent advances, Springer London, 2007.
Danielsson, K. , S. Selander , Multimodal texts in disciplinary education: A comprehensive framework, Springer Switzerland, 2021.
Deegan, M. , K. Sutherland , Text editing, print and the digital world, Routledge, 2016.
Doerner, R. , W. Broll , P. Grimm & B. Jung (eds), Virtual and augmented reality (VR/AR), Springer Nature Switzerland, 2022.
Duckett, J. , HTML and CSS: Design and build websites, John Wiley & Sons, 2014
Dutoit, T. , An introduction to text to speech synthesis, Springer Netherlands, 2013.
Eisenstein, J. , Introduction to natural language processing, MIT Press, 2019
Flanagan, D. , JavaScript: the definitive guide, O'Reilly Media, 2017.
Frain, B. , Responsive Web design with HTML5 and CSS, Packt Publishing Ltd., 2022.
Gillam, R. , Unicode demystified: A practical programmer's guide to the encoding standard, Addison-Wesley Professional, 2003.
Godse, A. P. , D A Godse , Multimedia technologies", Technical Publications, 2020.
Goodman, D. , Dynamic HTML – the definitive reference, O'Reilly Media, 2007
Iglezakis, I. , T.-E. Synodinou (eds), E-publishing and digital libraries: Legal and organizational issues, IGI Global, 2010.

Jurafsky, D. , J. H. Martin , Speech and language processing, Pearson Education India, 2013.
Prettyman, S. , Learn PHP 7: object oriented modular programming using HTML5, CSS3, JavaScript, XML, JSON and MySQL, Apress, 2016.
Robbins, J. N. , Learning Web design: A beginner's guide to HTML, CSS, JavScript and Web graphics, O'Reilly Media, 2012.
Royds, A. B. , K. Cagle & D. Storey , Using SVG with CSS3 and HTML5, O'Reilly Media, 2018.
White, B. , HTML and the art of authoring for the World Wide Web, Springer US, 2012.
Zaramenskikh, E. , A. Fedorova , Digital transformation and new challenges, Springer Nature Switzerland, 2020.

Image

Andina, T. , T. Dreier (eds), Digital ethics: The issue of images, Nomos, 2022.
Armi, L. , S. Fekri-Ershad , Texture Image Analysis and Texture Classification Methods: A review", 2019. https://arxiv.org/ftp/arxiv/papers/1904/1904.06554.pdf
Bai, M. , R. Urtasun , Deep Watershed Transform for Instance Segmentation, 2017
Bay, H. , T. Tuytelaars & L.V. Gool , SURF: Speeded up robust features, Computer Vision and Image Understanding, vol. 110, no. 3, pp. 346–359, 2008.
Bengio, Y. , Learning deep architectures for AI, Foundations and Trends in Machine Learning, vol. 2, no. 1, pp. 1–127, 2009.
Berg, M. , J. Boyar , L. M. Favrholdt & K. M. Larsen , Online minimum spanning trees with weight predictions, 2023. https://arxiv.org/abs/2302.12029
Breiman, L. , Random forests, Machine Learning, vol. 45, no. 1, pp. 5–32, 2001.
Burger, W. , M. J. Burge , Principles of digital image processing: Advanced methods, Springer Verlag, 2013.
Cai, Z. , N. Vasconcelos , Cascade R-CNN: Delving into high quality object detection, 2017. https://arxiv.org/abs/1712.00726
Chen, L.-C. , G. Papandreou , K. Murphy & A. L. Yuille , DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution and fully connected CRFs, 2017. https://arxiv.org/abs/1606.00915
Christianini, N. , & J. C. Shawe-Taylor , An introduction to support vector machines and other kernel based learning methods, Cambridge University Press, 2000.
Coppersmith, D. , S. J. Hong & J. R. M. Hosking , Partitioning nominal attributes in decision trees, Data Mining and Knowledge Discovery, vol. 3, pp. 197–217, 1999.
Csurka, G. , C. R. Dance , L. Fan , J. Williamowski & C. Bray , *Visual categorization with bags of key-points*, Proceedings of ECCV International Workshop on Statistical Learning in Computer Vision, 2004.
Cyganek, B. , Object detection and recognition in digital images, Wiley, 2013.
Das, R. , Content-based image classification, CRC Press, 2020.
Dryden, I. L. , K. V. Mardia , Statistical shape analysis, Wiley, 2016.
Ester, M , Kriegel, H.-P. , Sander, J. , Xu, X. , Simoudis, E. , Han, J. & Fayyad, U. M. (eds.). A density-based algorithm for discovering clusters in large spatial databases with noise. Proceedings of the Second International Conference on Knowledge Discovery and Data Mining (KDD-96), AAAI Press. pp. 226–231, 1996.
Fortner, B. , T. E. Meyer , Number by colors, Springer, 1997.
Girshick, R. , J. Donahue , T. Darrell & J. Malik , Region based convolutional networks for accurate object detection and segmentation, IEEE Transactions of Pattern Analysis and Machine Intelligence, 2015.
Gonzalez, R. C. , R. E. Woods , Digital image processing, Prentice Hall, 2008.
Guild, J. , The colorimetric properties of the spectrum, Pholosophical Transactions, vol. 230, no. 685, pp. 149–187, 1931.
Gunturk, B. K. , X. Li (eds), Image restoration: fundamentals and advances, CRC Press, 2018.
Haar, A. , "Zur Theorie der orthogonalen Funktionensysteme", Mathematische Annalen, vol. 69, pp. 331–371, 1910.
Haralick, R.M. , K. Shanmugan & I. Dinstein , Textural features for image classification, IEEE Transactions on Systems, Man, and Cybernetics, vol. SMC-3, pp. 610–621, 1973.
Hastie, T. , R. Tibshirani , & J. Friedman , The elements of statistical learning, Second edition. Springer, 2008.
Haykin, S. , Neural networks and learning machines, Third edition, Prentice Hall, 2009.
He, K. , G. Gkioxari , P. Dollar & R. Girshick , Mask R-CNN, 2018. https://arxiv.org/abs/1703.06870
Hung C-C , E. Song , Y. Lan , Image texture analysis: Foundations, models and algorithms, Springer, 2019.
Hunt, R. W. G. , D. Wright , in Luo, R. (ed.), Encyclopedia of color science and technology, Springer, 2014.
Hunter, R. , Photoelectric color-difference meter, Journal of Optical Society of America, vol. 38, no. 7, p. 661, 1948.
Jain, A. K. , Fundamentals of digital image processing, Prentice Hall, p. 439, 1989.
Kallen, S. A. , Digital cameras and camcorders, Lucent Books USA, 2014.
Kriss M (ed.), Handbook of digital imaging, Wiley, 2015.
Lancaric, S. , A survey of shape analysis techniques, Pattern Recognition, vol. 31, no. 8, pp. 983–1001, 1998.
Lei, T. , A. K. Nandi , Image segmentation: principles techniques and applications, Wiley, 2022.
Liu, C. Z. , S. Ramakrishnan , Image recognition: Progress trends and challenges, Nova Science Publishers, 2020.
Liu, W. , et al. , SSD: Single shot multibox detector, 2016.
Long, J. , E. Shelhamer , T. Darrell , Fully convolutional networks for semantic segmentation, 2015. https://arxiv.org/abs/1605.06211
Lowe, D. G. , Distinctive image features from scale-invariant key-points, International Journal of Computer Vision, vol. 60, no. 2, pp. 91–110, 2004.
Lucy, L.B. , "An iterative technique for the rectification of observed distributions", Astronomical Journal, vol. 79, no. 1, pp. 745–754, 1974.
MacQueen, J. B. Some methods for classification and analysis of multivariate observations, Proceedings of 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1. University of California Press. pp. 281–297, 1967. https://projecteuclid.org/euclid.bsmsp/1200512992

Mallat, S. G. , A theory for multiresolution signal decomposition: The wavelet representation, IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 11, no. 7, pp. 674–693, July 1989.
Michael, W. , Color calibration techniques for true color measurement, Lap Lambert Academic, 2015.
Otsu, N. , A threshold selection method from gray-level histograms, IEEE Transactions on Systems, Man, and Cybernetics, vol. 9, no. 1, pp. 62–66, 1979.
Poynton, C. , A guided tour of color space, Proceedings of the SMPTE Advanced Television and Electronic Imaging Conference, pp. 167–180, 1995.
Rajput, S. S. , N. U. Khan , A. K. Singh & K. V. Arya (eds.), Digital image enhancement and reconstruction, Elsevier Science, 2022.
Reed, C. S. , Copyright workflow for photographers, Peachpit Press, 2015.
Richardson, W. , Bayesian based iterative method of image restoration, Journal of the Optical Society of America (JOSA), vol. 62, no. 1, pp. 55–59, 1972.
Roussopoulos, N. , S. Kelley , F.D.R. Vincent , “Nearest neighbor queries”, Proceedings of 1995 ACMSIGMOD International Conference on Management of Data – SIGMOD ‘95, p. 71, 1995.
Sinha, P. K. , Image acquisition and preprocessing for machine vision systems, SPIE, 2012.
Soydaner D , Attention mechanisms in neural networks: Where it comes and where it goes, 2022. https://arxiv.org/abs/2204.13154
Strano, M. , B. M. Colosimo , Logistic regression analysis for experimental determination of forming limit diagrams, International Journal of Machine Tools and Manufacture, vol. 46, no. 6, pp. 673–682, 2006.
Sundararajan, D. , Digital image processing: a signal processing and algorithmic approach, Springer Nature Singapore, 2017.
Tan, M , R. Pang , Q. V. Le , EfficientDet: scalable and efficient object detection, 2020. https://arxiv.org/abs/1911.09070
Wiener, N. , Extrapolation interpolation and smoothing of stationary time series, John Wiley & Sons Inc., 1949.
Wright, W. D. , A re-determination of the trichromatic coefficients of the spectral colours, Transactions of the Optical Society, vol. 30, no. 4, pp. 141–164, 1929
Xia, F. , J. Liu , H. Nie , Y. Fu , L. Wan & X. Kong , Random walks, a review of algorithms and applications, 2020. https://arxiv.org/abs/2008.03639
Young, T. , On the theory of light and colors, Philosophical Transactions of the Royal Society of London, vol. 92, pp. 12–48, 1802.
Zhou, Z. , M. R. Siddiquee Md , N. Tajbakhsh , & J. Liang , UNet++: A nested U-net architecture for medical image segmentation, 2018. https://arxiv.org/abs/1807.10165

Graphics

Chakraborty, S. , Fundamentals of computer graphics, Everest Publishing House, 2010.
Boor, C. , Subroutine package for calculating with B-splines, Tech Rep. LA-4728-MS, Los Alamos Sci. Lab, Los Alamos, NM, pp. 109–121, 1971.
Foley, J. D. , A. van Dam , S. K. Feiner , et al. 1995. Computer graphics: Principles and practice. Addison-Wesley Professional, 1990.
Godse, A. P. , D. A. Godse , Computer graphics and multimedia, Technical Publications, 2021.
Hearn, D. & M. P. Baker , Computer graphics, C version 2nd edition. Prentice Hall, 1996.
Kane, E. M. , Ethics: A graphic designer's field guide, Simon & Schuster, 2010
Mandelbrot, B. B. , The fractal geometry of Nature, W H Freeman and Co., 1982.
Marsh, D. , Applied geometry for computer graphics and CAD, Springer, 2005.
Mathews, J. H. & C. D. Fink , Numerical methods using MATLAB, Prentice Hall, 2004.
O'Rourke, M. , Principles of three dimensional computer animation, W. W. Norton & Company, 2003.
Olive, J. , Maths: A student's survival guide, Cambridge University Press, 2003.
Rovenski, V. , Modeling of curves and surfaces with MATLAB. Springer, 2010.
Shirley, P. , Fundamentals of computer graphics. AK Peters/CRC Press, 2002.

Audio

Cardone, E.P. , The art and science of surround and stereo recording, Springer Vienna, 2021.
Cobussen, M. , N. Nielsen , Music and ethics, Routledge, 2016.
Davies, M.J. , Legal issues in the music industry, BuzzGig LLC, 2009.
Fourier, J.B.J. , Theorie analytique de la chaleur, Chez Firmin Didot, 1822.
Huber, D.M. , R. Runstein , Modern recording techniques, Routledge, 2017.
Jackson, W. , Digital audio editing fundamentals, APress, 2015.
Kim, H. , L. Remaggi , P. Jackson & A. Hilton , “Immersive spatial audio reproduction for VR/AR using room acoustic modelling from 360° images”, IEEE Conference on Virtual Reality and 3D User Interfaces (VR), pp. 120–126, 2019.
Olson H. F. , “Acoustical Engineering”, Van Nostrand, Princeton, N.J., 1957.
Pohlmann, K. , Principles of digital audio, McGraw-Hill, 2010.
Rossing, T. , R. Moore & P. Wheeler , Science of sound, Pearson, 2011.
Rumsey, F. , Desktop audio technology: Digital audio and MIDI principles, Routledge, 2003.
Schutze, S. , A.I. Schutze , New realities in audio: A practical guide for VR, AR, MR and 360 video, CRC Press, 2018.
Smith, J.O. , Introduction to digital filters: With audio applications, Createspace Independent Publishers, 2006.
Somberg, G. (editor), Game audio programming: Principles and practices, CRC Press, 2010.
Toole, F.E. , Sound reproduction: The acoustics and psychoacoustics of loudspeakers and rooms, Routledge, 2017.

Video

Bradski, G. R. , Computer vision face tracking for use in a perceptual user interface, IEEE Workshop on Applications of Computer Vision, pp. 214–219, 1998.

Comaniciu, D. , P. Meer , Mean Shift: A robust approach toward feature space analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence (PAMI), vol. 24, no. 5, pp. 603–619, 2002.

Crow, F. , Summed up tables for texture mapping, Proceedings of the SIGGRAPH, vol. 18, no. 3, pp. 207–212, 1984.

Dalal, N. & B. Triggs . Histograms of oriented gradients for human detection, IEEE Computer Society Conference on Computer Vision and Pattern Recognition, vol. 1, pp. 886–893, 2005.

Dasiopoulou, S. , et al. , Knowledge-assisted semantic video object detection, IEEE Transactions on Circuits and Systems for Video Technology, vol. 15, no. 10, pp. 1210–1224, 2005.

Farneback, G. , “Two-frame motion estimation based on polynomial expansion”, Proceedings of the 13th Scandinavian Conference on Image Analysis, 2003.

Haar, A. , Zur Theorie der orthogonalen Funktionensysteme, Mathematische Annalen, vol. 69, pp. 331–371, 1910

Horn, B. K. P. & B. G. Schunck , Determining optical flow, Artificial Intelligence, vol. 17, pp. 185–203, 1981

Jackson, W. , Digital video editing fundamentals, Apress, 2016.

Kallen, S. A. , Digital cameras and camcorders, Lucent Books, 2014.

Leggett, E. R. , Digitization and digital archiving, Rowman & Littlefield, 2020.

Lucas, B. D. and Kanade, T. , An iterative image registration technique with an application to stereo vision, Proceedings of Imaging Understanding Workshop, pp. 121–130, 1981.

Musburger, R. B. , Single-camera video production, CRC Press, 2012.

Poynton, C. , “A Technical Introduction to Digital Video”, John Wiley & Sons Inc., p. 175, 1996.

Slater, J. , Modern television systems, CRC Press, 1991.

Taskiran, M. N. , F. Pinarbasi , Multidisciplinary approaches to ethics in the digital era, IGI Global, 2021.

Viola, P. , M. J. Jones , Rapid object detection using a boosted cascade of simple features, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), vol. 1, pp. 511–518, 2001.

Weynand, D. , M. Weise , How video works, CRC Press, 2012.

Zhang, S. , Single shot refinement neural network for object detection, Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4203–4212, 2018.

Animation

Akenine-Moller, T. , E. Haines & N. Hoffman , Real time rendering, CRC Press, 2008

Burr, S.L. , Entertainment Law in a nutshell, West Group, 2004

Chakraborty, S. Fundamentals of computer graphics, Everest Publishing House, 2010.

Foley, J.D. , A. van Dam , S.K. Feiner , et al. Computer graphics: Principles and practice. Addison-Wesley Professional, 1990.

Glassner, A.S. , Principles of digital image synthesis, Morgan Kauffman Publishers, 1995

Gouraud, H. , Continuous shading of curved surfaces, IEEE Transactions on Computers, vol. C-20, no. 6, pp. 623–629, 1971.

Hearn, D. , M.P. Baker . Computer Graphics, C Version Second edition, Prentice Hall, 1996.

O'Rourke, M. , Principles of three dimensional computer animation, W. W. Norton & Company, 2003.

Phong B T , Illumination for computer generated pictures, Communication of the ACM, vol. 18, no. 6, pp. 311–317, 1975.

Roth S D , Ray casting for modelling solids, Computer Graphics and Image Processing, Elsevier, vol. 18, no. 2, pp. 109–144, 1982.

Rovenski, V. , Modeling of curves and surfaces with MATLAB, Springer Germany, 2010.

Shirley, P. , R.K. Morley , Realistic ray tracing, A K Peters/CRC Press, 2/e, 2003.

Thomas, F. , O. Johnston , The illusion of life: Disney animation, Hyperion, 1981, reprint 1997.

Wilkins L , C. Painter , P. Patterson , Media ethics: Issues and cases, Rowman & Littlefield, 2021.

Zeman, N.B. , Essential skills for 3D modeling, rendering and animation, CRC Press, 2015.

Compression

Falconer, K. , Fractal geometry: Mathematical foundations and applications, John Wiley & Sons, 2003.

Liebchen, T. , T. Moriya , N. Harada & Y. Kamamoto , The MPEG-4 ALS standard: Technology and applications, AES Convention, USA, 2005. http://elvera.nue.tu-berlin.de/files/0737Liebchen2005.pdf

Marpe, D. , T. Wiegand & G.J. Sullivan , The H.264/MPEG-4 AVC standard and its applications, IEEE Communications Magazine, vol. 44, no. 10, pp. 134–143, 2005.

Puri, A. & A. Majumder , Guide to video compression: From sgms to HEVC, CRC Press, 2012.

Salomon, D. , Data compression, Third edition. Springer, 2004.

Shannon, C.E. , A mathematical theory of communication, The Bell System Technical Journal, vol. 27, pp. 379–423, 623–656, July, October, 1948.

Welch, T.A. , A technique for high-performance data compression, IEEE Computer, vol. 17, no. 6, 1984.

Ziv, J. , A. Lempel , A universal algorithm for sequential data compression, IEEE Transactions on Information Theory, vol. 23, no. 3, pp. 337–343, 1977.

Ziv, J. , A. Lempel , Compression of individual sequences via variable rate coding, IEEE Transactions on Information Theory, vol. 24, no. 5, pp. 530–536, 1978.