December 31, 2024 Vol. 6 No. 9 E-ISSN 3027-0952 P-ISSN 3027-2033

International Journal of

CONVERGENT AND INFORMATICS SCIENCE RESEARCH (IJCISR)

www.harvardpublications.com

EVALUATING THE QUALITY OF METHODS AND METRICS IN DIGITAL MULTIMEDIA SYSTEMS - A REVIEW

[1]KAREEM AFISS EMIOLA ADEBOWALE; & [2]AWODELE SIMON OLUFIKAYO

[1]Department of Cybersecurity, University of Ilesa, Ilesa, Nigeria. [2]New Horizons Systems Solutions, No 5, Babatola Drive, Off Awolowo Way, Ikeja, Lagos, Nigeria.

Corresponding Author: *afisskareem@unilesa.edu.ng*

DOI: https://doi.org/10.70382/hijcisr.v06i9.005

Abstract

The rapid advancement of digital multimedia technologies has necessitated rigorous evaluation methodologies to ensure that they meet user expectations and standards. Digital multimedia systems have become integral to various applications ranging from entertainment to education and communication. The quality of these systems significantly impacts user experience and satisfaction. Service providers are enacting different quality of service (QoS) solutions to issue the best quality of experience (QoE) to their customers.

Keywords: Quality Assessment, Objective Quality Metric, Subjective Quality Metric, Multimedia Quality, Audiovisual Perception, Quality of Service

Thus, devising precise

INTRODUCTION

In the rapidly evolving landscape of digital multimedia systems, the demand for high-quality content, efficient processes, and robust performance evaluations has become increasingly paramount. As these systems serve diverse applications-from streaming services and online gaming to virtual reality and augmented reality-the need for a systematic approach to assess their quality has never been more critical. Digital multimedia systems encompass a broad range of technologies that integrate various forms of content, including text, graphics, audio, and video, into unified experiences (Chalvatzaki & Tziritas, 2022). These systems are characterized by their ability to deliver rich, interactive content to users through different platforms and devices. The complexity of multimedia content, coupled with varying user expectations and contextual factors, introduces significant challenges in evaluating

quality metrics that will greatly help to improve multimedia services over networks. In this paper, we provide a comprehensive analysis of the works that have been carried out over recent years in audio, video and audiovisual quality assessments, describing existing methodologies in terms of requirement of a reference signal, feature extraction, feature mapping and classification schemes. The review also aims to explore the existing methods and metrics used to assess the quality of multimedia systems, considering both subjective and objective approaches. We will provide a comprehensive evaluation of their effectiveness, applicability, and limitations, thereby identifying areas for future research.

the quality of these systems.

Evaluating the quality of digital multimedia systems involves assessing both the content being delivered and the methods employed in its processing and transmission. The quality perceived by users can be influenced by numerous factors, including playback performance, visual and audio fidelity, interactivity, and user experience (Ali & Hameed, 2023). Thus, the need for effective quality evaluation methods and metrics becomes essential to ensure that these systems meet the expected standards of excellence.

Quality assessments can be classified into two broad categories: objective evaluation and subjective evaluation (Ali & Hameed, 2023). Objective evaluation typically relies on quantitative measurements that can be automated and standardized, such as bit rate, frame rate, and signal-to-noise ratio. In contrast, subjective evaluation engages human participants to assess quality based on their experience, often through surveys or psychophysical tests. Each approach has its merits and challenges, and a comprehensive evaluation framework must balance both perspectives.

OVERVIEW OF MULTIMEDIA QUALITY ASSESSMENT

Multimedia quality assessment (MQA) refers to the systematic evaluation of the quality of multimedia content, which includes audio, video, and images (Beerends et al., 2016). Given the ubiquity of multimedia across various platforms, such as streaming services, gaming, social media, and virtual reality, ensuring high-quality delivery is essential for user satisfaction, engagement, and retention. Effectively assessing quality is critical for developers, content creators, and service providers to understand how well their products perform and what improvements may be needed.

Multimedia quality can be analyzed through multiple dimensions, each contributing to the overall user experience (Chen et al., 2023; Fan et al., 2022):

1. Technical Quality: This pertains to the fidelity and performance of the multimedia content. It includes parameters such as resolution, frame rate, bit rate, and audio clarity. Technical quality is crucial for ensuring that the content is rendered as intended without artifacts or degradation.

2. Perceived Quality: This is subjective and varies based on individual user experience. It encompasses factors such as aesthetic appeal, user interface design, and the emotional

impact of the content. Perceived quality directly influences user satisfaction and engagement.

3. User Experience (UX): User experience extends beyond technical specifications to incorporate usability, accessibility, and the overall flow of interactions within the multimedia content. A seamless UX is essential for keeping users engaged and satisfied.

4. Contextual Quality: The environment in which multimedia is consumed can affect its perceived quality. Factors like network conditions, device capabilities, and user context (e.g., user preferences and stress levels) can impact how content is interpreted and enjoyed.

Developing Effective Methods and Metrics

The development of methods and metrics for quality evaluation is a dynamic and multifaceted process (Gomez et al., 2023). Key considerations in this domain include:

i. Relevance: Metrics must be relevant to the specific multimedia context being evaluated. For example, metrics for video quality may differ significantly from those appropriate for audio quality or interactive systems.

ii. Standardization: The establishment of universal standards and benchmarks can facilitate consistent quality assessments across different systems and applications. Notable standards include the International Telecommunication Union (ITU) recommendations and the Moving Picture Experts Group (MPEG) standards.

iii. Usability: The chosen methods and metrics should be user-friendly and applicable in real-world scenarios. Complex evaluation procedures may be impractical for widespread use in industry settings.

iv. Integration: Effective evaluation should integrate various quality dimensions, including technical attributes, user experience, and accessibility considerations. A multi-dimensional approach ensures a holistic understanding of quality.

v. Adaptability: Given the rapid technological advancements in multimedia, evaluation methods and metrics must evolve to address emerging technologies, such as AI-driven content generation and immersive experiences in virtual and augmented reality.

Methods of Quality Assessment

Multimedia quality assessment can be categorized into two primary approaches (Hines et al., 2015; Zhang et al., 2023): objective and subjective evaluations.

i. Objective Evaluation: This approach employs mathematical algorithms and models to quantify quality based on measurable parameters. Some key objective metrics include:

- Peak Signal to Noise Ratio (PSNR): Often used for images and videos, it compares the maximum possible power of a signal to the power of corrupting noise.
- Mean Opinion Score (MOS): A metric derived from user ratings, in which users score the quality of content on a predefined scale (e.g., 1 to 5).

- Structural Similarity Index (SSIM): This evaluates image quality based on structural information, comparing luminance, contrast, and structure.

ii. Subjective Evaluation: Involves human participants providing feedback based on their experience with the multimedia content. Common methods include:
   - Quality Assessment Surveys: Using scales to capture perceived quality from a user's perspective.
   - Psychophysical Tests: Experimental setups where users evaluate multiple samples for preferences or noticeable differences.

Challenges in Multimedia Quality Assessment

Despite advancements in assessment techniques, several challenges persist (Zhou et al., 2014):

i. Subjectivity and Variability: Different users may have varying perceptions of quality based on personal preferences, cultural backgrounds, and experiences. This subjectivity can lead to inconsistent results in assessments.

ii. Dynamic Content: Multimedia content is often dynamic, involving real-time changes (e.g., live streaming, interactive environments). Assessing quality in these contexts can be complex and requires adaptive metrics.

iii. Technological Advancements: The rapid pace of technology, such as advancements in codecs, streaming protocols, and display technologies, means that existing metrics and models may quickly become outdated or less relevant.

iv. Context Dependency: Quality perception can heavily depend on contextual factors, such as the viewing environment (light conditions, screen size) and the device used (smartphone versus large screen). Creating universal metrics that account for all contexts is challenging.

Future Directions in Multimedia Quality Assessment

The evolution of multimedia technology will drive ongoing research in quality assessment methodologies (Picardi & Mazzola, 2023). Emerging areas of focus include:

- Artificial Intelligence and Machine Learning: As these technologies become more sophisticated, they can be employed to predict quality based on existing data and enhance automatic assessments. AI can also help tailor content delivery based on individual user profiles.
- Real-Time Assessments: Developing metrics that can provide real-time quality feedback during content delivery will help stakeholders address issues proactively.
- User-Centric Metrics: Shifting towards user-centric evaluation frameworks that consider context and individual preferences can improve quality assessments and enhance user satisfaction.
- Integration of Quality Dimensions: A comprehensive approach that integrates technical, perceived, UX, and contextual quality into unified frameworks may offer a holistic view of multimedia quality.

FACTORS INFLUENCING AUDIOVISUAL MULTIMEDIA QUALITY

Audiovisual multimedia quality encompasses the quality of both audio and video components in multimedia content. Several factors influence this quality, affecting users' overall experience and satisfaction (Zahid & Tiago, 2017). Understanding these factors is crucial for developers, content creators, and service providers to enhance the quality of audiovisual content effectively.

i. Technical Factors

- Bitrate: A higher bitrate often results in better audio and video quality, as it allows for more data to be transmitted per second. Low bitrates can lead to compression artifacts, such as blockiness in video and muddy sound in audio.
- Resolution: The resolution defines the clarity of the video. Higher resolutions (e.g., 1080p, 4K) provide more detailed images. However, the ability of a device to render high resolution must be considered.
- Frame Rate: The number of frames displayed per second (e.g., 30fps, 60fps). Higher frame rates can create smoother motion, which is particularly important for fast-paced content like sports or action movies.
- Audio Sample Rate and Bit Depth: The sample rate (e.g., 44.1 kHz, 48 kHz) affects audio clarity and fidelity. Higher sample rates can capture more detail. Bit depth (e.g., 16-bit, 24-bit) influences dynamic range and noise levels in audio.
- Compression: The technique used to compress audio and video affects quality. Lossy compression (e.g., MP3, MPEG) sacrifices some quality for smaller file sizes, while lossless compression (e.g., FLAC, MKV) maintains quality but results in larger files.

ii. Content Factors

- Source Material Quality: The initial quality of the content, such as resolution and fidelity of the source (e.g., film, broadcast), significantly impacts the final output. Poor source material leads to compromised quality, regardless of the technology used for distribution.
- Encoding and Decoding: The methods and algorithms used for encoding (compressing and preparing) and decoding (playing back) audiovisual content can affect perceived quality. High-quality encoding settings can preserve more detail.
- Genre and Format: Different genres may require different quality characteristics. For example, documentaries may prioritize clear dialogue over background music, while action films may focus on high dynamic range and color.

iii. Delivery Factors

- Network Conditions: Bandwidth and network stability play significant roles in streaming quality. Low bandwidth can lead to buffering, lower resolution streaming, and compression artifacts. Latency can also affect real-time communications (e.g., video calls).

- Streaming Protocols: The choice of protocols (e.g., HTTP Live Streaming, MPEG-DASH) influences how well the content adapts to varying network conditions, affecting quality.
- Device Compatibility: The capabilities of the device (e.g., screen size, hardware capability) affect how audiovisual content is rendered. Older devices may struggle with high-resolution or bitrate content, leading to poor quality.

v. Environmental Factors
- Viewing Environment: Ambient lighting, screen size, and viewing distance can affect the perceived quality of video. A well-lit environment may wash out colors on a screen that's not sufficiently bright.
- Audio Environment: Background noise and acoustics of the environment can significantly influence audio quality. Listening through headphones in a quiet space differs from experiencing sound in a noisy room.

vi. User Factors
- User Preferences and Expectations: Individual tastes can heavily affect perceived quality. Some users prioritize high-resolution video for aesthetics, whereas others may prioritize clear dialogue.
- Fatigue and Attention: User engagement and fatigue levels may influence how audiovisual content is perceived. Users may not notice lower quality when they are highly engaged or when experiencing fatigue.
- Cognitive Load: The ability to process audiovisual information can vary among individuals. Excessively complicated visuals or audio details can overwhelm viewers, leading to decreased perceived quality.

vii. Perceptual Factors
- Visual Perception: Human visual perception is affected by factors such as color contrast, brightness, and motion. Content must be carefully crafted to maximize the effectiveness of these visual cues.
- Audio Perception: The human ear is sensitive to frequency ranges, and audio should be mixed to maintain clarity across these frequencies. Elements like balance among speech, music, and sound effects are crucial for overall quality.

CAUSES OF DEGRADATION OF AUDIOVISUAL SIGNALS

The degradation of audio and visual signals can occur due to a variety of factors, both during the capture and transmission processes. Understanding these causes is critical for improving the quality of audiovisual content and ensuring a satisfactory user experience (Mamun et al., 2015; Moller & Raake, 2014).

Causes of Degradation of Audio Signals

1. Noise Interference:
   - Electronic Interference: Background noise created by electronic devices (e.g., motors, fluorescent lights) can introduce hum and buzz into audio recordings.

- Acoustic Noise: Environmental sounds such as traffic, crowd noise, or wind can be captured unintentionally, degrading the quality of the intended audio signal.

2. Low Bitrate and Compression:
   - Lossy Compression: Techniques like MP3 or AAC reduce file sizes by removing audio data, which can lead to loss of detail, artifacts, and a reduction in dynamic range and clarity.
   - Bitrate Limitations: Audio encoded at a low bitrate may result in a flat or muddy sound.
3. Poor Microphone Quality:
   - Inadequate Equipment: Low-quality microphones may not capture the full range of audio frequencies accurately, leading to a dull or distorted sound.
   - Placement Issues: Improper positioning of microphones can result in unwanted sound pickup or loss of desired sound.
4. Room Acoustics:
   - Reverberation: Excessive echo or reverb in a recording environment can muddy the sound and obscure clarity.
   - Reflection and Absorption: Hard surfaces can reflect sound waves, creating troublesome echoes, while soft surfaces may absorb too much sound, leading to a lack of presence.
5. Signal Distortion:
   - Clipping: When audio signals exceed the maximum amplitude that the capture equipment can handle, clipping can occur, resulting in harshness and distortion.
   - Harmonic Distortion: Alteration of the original audio waveform during processing can lead to unwanted tonal changes.
6. Sampling Rate Issues:
   - Inadequate Sampling Rate: If the sampling rate is too low, high-frequency sounds may be improperly captured, leading to aliasing and loss of audio detail.

Causes of Degradation of Visual Signals

1. Compression Artifacts:
   - Lossy Compression: Video formats that use lossy compression (e.g., H.264) may exhibit artifacts, such as blockiness, blurring, and color banding, particularly at low bitrates.
   - Over-Compression: Excessive compression can significantly degrade image quality, leading to loss of detail and fidelity.
2. Low Resolution:
   - Insufficient Resolution: Content captured or transmitted at low resolutions (e.g., VGA instead of HD) will appear pixelated and lack detail, especially when viewed on larger displays.
3. Poor Lighting Conditions:
   - Underexposure: Insufficient light can lead to noisy images, poor color representation, and details being lost in shadows.

- Overexposure: Too much light can wash out details, causing highlights to be blown out and color fidelity to suffer.

4. Camera Movement:
   - Shaky Footage: Unstable shooting can lead to motion blur and lack of focus, making visual content difficult to watch.
   - Focus Issues: Improper focus can result in scenes that are out of focus, degrading overall clarity.
5. Signal Interference:
   - Electromagnetic Interference: Various electronic and radio frequency devices can interfere with video signals, causing artifacts, static, or distortion.
   - Transmission Errors: During transmission over networks or broadcast, signal loss can corrupt the video stream, leading to dropped frames or pixelation.
6. Display Issues:
   - Screen Quality: The quality of the display (e.g., resolution, panel technology) affects perceived video quality. Low-quality displays can distort colors, contrast, and sharpness.
   - Color Calibration: Poorly calibrated displays may misrepresent colors, leading to an inaccurate and unsatisfactory visual experience.

Causes of Degradation for Both Audio and Visual Signals

1. Bandwidth Limitations: Insufficient bandwidth during streaming can result in both audio and video signals being compressed more than necessary, leading to degradation in quality.
2. Encoding/Decoding Errors: Errors that occur during the encoding or decoding process can introduce visual glitches, artifacts, and audio dropouts.
4. Synchronization Issues: Delays in audio and visuals not being properly aligned (lip-sync issues) can degrade the experience, even if the audio and video are of good quality individually.

QUALITY OF SERVICE, QUALITY OF EXPERIENCE AND QUALITY OF PERCEPTION

Three significant aspects are Quality of Service (QoS), Quality of Experience (QoE), and Quality of Perception (QoP). Each plays a distinct role in the overall success and functionality of multimedia systems (Yan & Mou, 2016).

Quality of Service (QoS): QoS in multimedia systems refers to the technical measures that ensure optimal performance, reliability, and efficiency of the system delivering multimedia content (Zahid & Tiago, 2017). It involves quantifiable parameters critical for maintaining high-quality user experiences.

i. Key Metrics
   - Bandwidth: The amount of data that can be transmitted over a network within a given time period. Higher bandwidth is essential for smooth video streaming and high-quality audio playback.

- Latency: The delay time between sending a request and receiving the response. In multimedia applications (like video conferencing), lower latency is crucial to prevent lag and ensure seamless interaction.
- Packet Loss: The percentage of packets that fail to reach their destination. In multimedia streaming, packet loss can result in interruptions like freezes or distortion in audio and video.
- Jitter: The variation in packet arrival time. High jitter can lead to uneven video playback and choppy audio, significantly impacting user experience.
- Error Rate: The frequency of errors that occur during data transmission. High error rates can degrade the perceived quality of multimedia services.

ii. Importance in Multimedia Systems

- Service Scalability: QoS guarantees that multimedia services can scale to accommodate many users without degrading quality, crucial for platforms like video streaming services.
- User Satisfaction: By ensuring minimal latency, high bandwidth, and low packet loss, QoS directly enhances user satisfaction, which is vital for retention in competitive multimedia markets.
- Real-Time Services: For applications that require real-time interaction (e.g., video conferencing, online gaming), ensuring QoS is essential to maintaining a smooth experience.

iii. Implementation Techniques

- Traffic Prioritization: Utilizing methods like Differentiated Services (DiffServ) to prioritize multimedia traffic over less time-sensitive data (e.g., email).
- Resource Reservation Protocols: Techniques like RSVP (Resource Reservation Protocol) allow applications to reserve network resources in advance to guarantee QoS.
- Multi-tiered Architecture: Designing systems that separate control and data planes can help optimize the delivery of multimedia content under variable network conditions.

Quality of Experience (QoE): QoE goes beyond the technical metrics of service quality to encompass the overall satisfaction and perception of the user while consuming multimedia content. It is a multidimensional view that includes emotional, cognitive, and contextual factors affecting users' experiences (Liu et al., 2022).

i. Key Components

- Content Quality: The resolution, bitrate, and overall technical quality of the multimedia content itself. For video, this includes resolution (e.g., 4K vs. 1080p), frame rate, and audio fidelity.
- User Interface (UI) Design: The usability and aesthetics of the interface through which users interact with multimedia content. A well-designed UI enhances engagement and satisfaction.
- Personalization: Adaptation of content and recommendations based on user preferences, which can vastly improve satisfaction by providing relevant content.

- Viewer Context: External factors such as the user's environment, the device used (smartphone, tablet, PC), and concurrent tasks can influence QoE.

ii. Importance in Multimedia Systems

- Customer Retention: A high QoE leads to user loyalty. When audiences feel that their experience is enjoyable, they are more likely to return to the platform or service.
- Competitive Advantage: Multimedia platforms that prioritize QoE through seamless delivery, engaging interfaces, and tailored content can differentiate themselves in competing markets.
- Market Research and Development: Understanding QoE allows developers to fine-tune multimedia applications, focusing on user feedback for improvement.

iii. Assessment Methods

- Subjective Evaluations: Surveys, focus groups, and user feedback mechanisms can gauge feelings about multimedia quality and satisfaction.
- Objective Measurements: Analysis of streaming metrics (buffering incidents, view durations) to correlate technical performance with perceived user satisfaction.

Quality of Perception (QoP): QoP refers to individual users' interpretations and emotional evaluations of multimedia content, influenced by past experiences, context, and subjective expectations. QoP is about the perception of quality rather than quantifiable metrics (Li et al., 2016; Dagiuklas, 2015).

i. Key Components

- Emotional Impact: The emotional response elicited by multimedia content (e.g., joy from a good film, frustration from buffering) significantly influences QoP.
- Cognitive Biases: Users' prior experiences and biases affect how they rate quality. Familiarity with a brand or content type can skew perceptions.
- Social Influence: Peer opinions, reviews, and social media discussions surrounding multimedia content can shape an individual's perception before they even engage with it.
- Contextual Relevance: How well the multimedia content resonates with the current mood, interests, and needs of the user significantly influences their perception of its quality.

ii. Importance in Multimedia Systems

- Brand Loyalty: Positive QoP can lead to strong brand attachment, where users continuously choose a specific multimedia service over competitors due to perceived quality.
- Content Creation: Understanding QoP can guide content creators in developing engaging and impactful multimedia products that resonate with audiences emotionally.
- User Engagement: Recognizing that perception significantly impacts engagement helps designers create multimedia experiences that captivate users' attention.

iii. Evaluation Techniques

- Qualitative Research: In-depth interviews and observation studies that explore how users interpret and react to multimedia content.
- Quantitative Surveys: Large-scale surveys examining user perceptions and emotional responses to specific content or services.

OVERVIEW OF EXISTING EVALUATION METHODS AND METRICS

A comprehensive overview of perceptual (objective) quality assessment methods for audio, video and audiovisual multimedia signals are presented here (Chen et al., 2023). Each signal type (i.e., audio, video, and audiovisual) is addressed in a different subsection.

A. **Methods and Metrics of Audio Quality Assessment:** Audio Quality Assessment (AQA) involves the evaluation of audio signals to determine their quality as perceived by human listeners. Assessing audio quality is essential for various applications, including broadcasting, streaming, music production, and telecommunications. These assessments can be categorized into subjective and objective methods, each employing different metrics and techniques.

- **Subjective Evaluation Methods**: Subjective methods rely on human listeners to evaluate audio quality based on personal perception. The assessments typically involve listening tests designed to gather qualitative and quantitative feedback. The methods are:

  i. Absolute Category Rating (ACR)
  - Listeners rate the overall quality of audio samples on a numerical scale (e.g., Mean Opinion Score - MOS) ranging from bad (1) to excellent (5) or even a 10-point scale.
  - ACR provides a direct insight into perceived audio quality.

  ii. Degradation Category Rating (DCR)
  - Participants compare a degraded audio sample with the original and categorize the perceived quality degradation.
  - This method is useful for quantifying the extent of quality loss in processed audio.

  iii. Paired Comparison (PC)
  - Listeners are given two audio samples and asked to indicate which one they prefer based on quality.
  - This method helps to discern small quality differences between samples.

  iv. Multiple Stimulus with Hidden Reference and Anchor (MUSHRA)
  - Listeners evaluate several audio samples, including one hidden reference (high-quality sample) and other lower-quality samples.
  - Participants rate these samples on a continuous scale against the reference, gathering detailed subjective quality assessments.

  v. Expert Listening Panels
  - Trained experts evaluate audio samples based on specific criteria, such as clarity, fidelity, spatial quality, and overall reproduction accuracy.
  - This method offers in-depth qualitative evaluations but requires trained personnel.

- **Objective Evaluation Methods:** Objective methods utilize computational techniques and algorithms to analyze audio signals and assess quality. These methods provide consistent, repeatable results and can be utilized for large-scale assessments (Picardi & Mazzola, 2023). Key Objective Methods are as follows:
  i. Signal-to-Noise Ratio (SNR)
  - SNR measures the ratio of the desired signal power to the background noise power.
  - Higher SNR values indicate a better audio quality, reflecting a clearer signal.

  ii. Perceptual Evaluation of Audio Quality (PEAQ)
  - Based on human auditory perception models, PEAQ predicts perceived audio quality by assessing various attributes, such as distortion and masking effects.
  - It outputs a score that correlates with listener assessments.

  iii. Weighted Signal-to-Noise Ratio (WSNR)
  - This metric accounts for the frequency sensitivity of human hearing by giving more weight to frequencies that are more critical for auditory perception.
  - WSNR provides a better correlation with perceived audio quality than traditional SNR.

  iv. Total Harmonic Distortion (THD)
  - Measures the amount of distortion present in an audio signal, quantified as a percentage of the original signal.
  - Lower THD values indicate better fidelity.

  v. Dynamic Range
  - The difference between the loudest and the quietest parts of an audio signal.
  - A larger dynamic range generally correlates with better audio quality, as it allows for both soft and loud sounds to be reproduced accurately.

  vi. Perceptual Loudness (e.g., LUFS)
  - Measures the perceived loudness of audio, capturing how loud a sound is to human ears, often quantified using Loudness Units Full Scale (LUFS).
  - Ensures consistent loudness levels across audio samples.

  vii. Columbia University Audio Quality Assessment (CQA)
  - A combination of various objective metrics designed to evaluate audio quality in a more comprehensive manner, accounting for multiple factors affecting human auditory perception.

- **Combined Metrics:** In addition to standard metrics, some approaches combine assessments for a more complete view of audio quality:
  i. Mean Opinion Score (MOS) for Objective Metrics
  - Integrates subjective ratings with objective measurements to provide a composite score that represents both human and computational evaluations.

ii. Multimedia Quality Assessment Index (MQAI)

- A combined index that integrates separate audio and video quality metrics into a single composite score, representing overall multimedia quality.

B. Methods and Metrics for Video Quality Assessment: Video Quality Assessment (VQA) is critical for evaluating the quality of video content, ensuring it meets the necessary standards for viewing experiences in applications such as streaming, broadcasting, and video production. Similar to audio quality assessment, video quality assessment can be classified into subjective and objective methods, each with distinct metrics and approaches (Zhang et al., 2023). Below is a comprehensive overview of the key methods and metrics used in Video Quality Assessment.

- **Subjective Evaluation Methods:** Subjective methods rely on human viewers for assessing video quality. These methods collect qualitative data based on personal perceptions of quality. Key Subjective Methods include:

  i. Absolute Category Rating (ACR)

  - Viewers rate the overall quality of video samples on a numerical scale (e.g., Mean Opinion Score - MOS) that commonly ranges from 1 (bad) to 5 (excellent) or a 10-point scale.
  - ACR provides direct feedback about users' perceptions of video quality.

  ii. Degradation Category Rating (DCR)

  - Participants view a degraded video sample and compare it against the original, categorizing the perceived quality degradation.
  - This helps quantify how much quality has been lost due to processing or compression.

  iii. Paired Comparison (PC)

  - Viewers are presented with two video samples and asked to indicate which one they prefer in terms of quality.
  - This method helps identify small differences in quality between two samples.

  iv. Multiple Stimulus with Hidden Reference and Anchor (MUSHRA)

  - Viewers evaluate multiple video samples, including one hidden reference and various degraded versions.
  - Participants give scores based on how they perceive the quality relative to the hidden reference, allowing for comprehensive assessments.

  v. Expert Panel Evaluation

  - A panel of trained experts evaluates videos based on specific criteria, including clarity, artifacts, motion smoothness, and overall aesthetic quality.
  - This method provides detailed qualitative assessments but can be resource-intensive.

- **Objective Evaluation Methods**: Objective methods utilize algorithms and computational techniques to analyze video signals and assess quality based on measurable attributes. These methods are precise, repeatable, and suitable for large datasets (Ali & Hameed, 2023). Key Objective Methods include:

i. Peak Signal-to-Noise Ratio (PSNR)
- Measures the ratio between the maximum possible signal power of a video to the noise power that affects its reproduction.
- Higher PSNR values indicate better video quality.

ii. Structural Similarity Index (SSIM)
- Compares the structural information between the original and distorted video frames. SSIM accounts for luminance, contrast, and structure in its calculations.
- SSIM scores range from 0 to 1, with 1 indicating perfect structural similarity.

iii. Video Multimethod Assessment Fusion (VMAF)
- Developed by Netflix, VMAF uses machine learning to correlate various quality metrics with human perceptions of video quality.
- It combines multiple objective metrics to provide a single composite score that aligns closely with viewer assessments.

iv. Mean Squared Error (MSE)
- This metric quantifies the average squared differences between pixel values in the original and distorted frames.
- Lower MSE values indicate better quality.

v. Temporal Quality Metrics (TQM)
- Metrics that evaluate the temporal coherence of video sequences, assessing aspects like blurriness during motion and frame rates.
- Important for determining the smoothness and fluidity of motion in videos.

vi. Bitrate and Compression Efficiency
- Assessing the bitrate at which video is encoded can indicate quality. Generally, a higher bitrate correlates with higher quality but requires more bandwidth.
- Compression metrics evaluate how much quality is retained versus the size decrease achieved.

vii. Noise and Artifact Detection
- Objective methods to measure noise levels, video artifacts (e.g., blocking, banding), and other issues that may affect perceived quality.
- Algorithms identify and quantify these artifacts, which can significantly reduce video quality.

- **Combined Metrics:** In addition to standalone measurements, some approaches combine various assessments to provide a more comprehensive view of video quality:

i. Mean Opinion Score (MOS) for Video
- A composite score aggregating subjective ratings from viewers and objective quality metrics to provide a unified assessment of video quality.

ii. Composite Quality Index (CQI)

- A metric that integrates various objective and subjective quality assessments into a single score, giving a holistic view of video quality.

C. **Methods and Metrics of Audiovisual Quality Assessment:** Audiovisual Quality Assessment (AQA) is essential in ensuring that both audio and video components of multimedia content deliver a satisfying user experience. It encompasses various methods and metrics to measure the quality of audio-visual content effectively (Liu et al., 2022). Below are the key methods and metrics for AQA, categorized into subjective and objective approaches.

- **Subjective Evaluation Methods:** Subjective methods rely on human listeners and viewers to assess audiovisual quality based on perception and experience (Gomez et al., 2023). Key subjective methods are:

i. Absolute Category Rating (ACR)

- Viewers or listeners rate the overall quality of audio-visual samples on a predefined scale (e.g., Mean Opinion Score - MOS).
- This method provides direct ratings of quality, typically from 1 (poor) to 5 (excellent).

ii. Degradation Category Rating (DCR)

- Participants compare a degraded audio-visual sample against the original version and categorize the perceived quality degradation.
- This approach helps quantify differences in perceived quality.

iii. Paired Comparison

- Subjects are presented with two audiovisual samples and asked to choose which one they prefer.
- This method facilitates side-by-side assessments to discern quality differences.

iv. Multiple Stimulus with Hidden Reference and Anchor (MUSHRA)

- Participants view multiple audio-visual samples, one being a hidden reference of best quality and others varying in quality.
- Viewers rate these samples relative to the reference, helping to gauge differences in quality perception.

v. Expert Listening/Viewing Panels

- Trained panels assess audio and video quality based on defined criteria such as clarity, spatial quality, and synchronization.
- This method provides in-depth evaluations but can be resource-intensive.

- **Objective Evaluation Methods:** Objective methods involve algorithms and models to evaluate audiovisual quality based on signal analysis, providing consistent and quick assessments (Li et al., 2016). This includes:

i. Video Quality Metrics

- Peak Signal-to-Noise Ratio (PSNR): Measures the ratio between the maximum possible signal power to noise power, commonly used for comparing original and encoded video. Higher PSNR indicates better quality.

- Structural Similarity Index (SSIM): Evaluates video quality by comparing structural information in original and processed frames. SSIM scores range from 0 to 1, with 1 indicating perfect quality.

- Video Multimethod Assessment Fusion (VMAF): Developed by Netflix, VMAF integrates various quality metrics using machine learning to predict perceived quality, providing a composite score that aligns closely with viewer perceptions.

ii. Audio Quality Metrics

- Signal-to-Noise Ratio (SNR): The ratio measuring signal power to noise power, indicating audio quality. Higher SNR values suggest better quality.

- Perceptual Evaluation of Audio Quality (PEAQ): A model simulating human auditory perception, producing scores that predict perceived audio quality.

- Loudness Measurement (e.g., LUFS): Assesses the perceived loudness of audio, helping ensure consistent volume levels across different media.

-Temporal and Synchronization Metrics

- Metrics that evaluate the temporal coherence and synchronization of audiovisual components, assessing aspects like lip-sync and dynamic motion in video alongside corresponding audio cues.

- **Combined Metrics for Audiovisual Quality Assessment:** In addition to separate evaluations of audio and video quality, some metrics combine these aspects for a comprehensive assessment.

i. Audiovisual Quality Index (AVQI)

- Integrates both audio and video quality scores, providing a single metric representing overall audiovisual quality.

ii. Mean Opinion Score (MOS) for A/V Synchronization

- A combination of subjective ratings for both audio and video quality, assessing how well the audiovisual components work together.

iii. Synchronization Error Metrics

- Measures the timing alignment of audio and video streams, quantifying how out of sync they are.

CONCLUSION

The evaluation of methods and metrics in digital multimedia systems is a crucial endeavor that significantly influences the design, development, and deployment of multimedia applications. This article provided a survey of existing multimedia quality evaluation methods with a focus on audio, video and audiovisual quality measurement techniques. The paper also presented a classification of audio and video quality metrics based on their underlying methodologies. This review has highlighted the diverse array of methodologies and criteria employed to assess multimedia quality, emphasizing the importance of a multi-faceted approach that encompasses both subjective and objective measures.

As digital multimedia continues to evolve, driven by advancements in technology and changing user expectations, the necessity for robust evaluation methods will only become more pronounced. It is essential that future research focuses on refining existing metrics, exploring new avenues for quality assessment, and fostering interdisciplinary collaboration to create a comprehensive evaluation ecosystem. Ultimately, a dedicated effort in evaluating and improving multimedia quality will pave the way for more engaging, efficient, and innovative applications that enhance user experiences across multiple platforms.

REFERENCES

Ali, R., & Hameed, A. (2023). A survey on quality of service and quality of experience in multimedia streaming. Multimedia Tools and Applications. 82(4), 5431-5455.

Beerends, J.G., Nieuwenhuizen, K., & Broek, E.L. (2016). Quantifying Sound Quality in Loudspeaker Reproduction. Journal of the Audio Engg. Society. vol. 64. no. 10.

Chalvatzaki, E., & Tziritas, G. (2022). Quality assessment of digital media: A review of metrics and methods. Applied Sciences. 12(6), 2023.

Chen, J., Wang, D., & Zhang, Q. (2023). An overview of recent advances in quality assessment metrics for multimedia content. Information Sciences. 623. 264-282.

Chen, C., Izadi, M., & Kokaram, A. (2016). A No-reference Perceptual Quality Metric for Videos Distorted by Spatially Correlated noise, ACM Multimedia

Dagiuklas, T. (2015). Multimedia quality of experience (QoE): current status and future requirements. Wiley.

Fan, J., Chen, O., & Yang, J. (2022). A comprehensive survey of 3D video quality assessment. Multimedia Tools and Applications. 81(20). 28427-28454.

Gomez, J., Arrebola, D., & Schmitt, R. (2023). Quality evaluation of interactive multimedia systems: Towards a user-centered approach. IEEE Transactions on Multimedia. 25. 1011-1021.

Hines, A., Gillen, E., Kelly, D., Skoglund, J., Kokaram, A., & Harte, N. (2015). ViSQOLAudio: An objective audio quality metric for low bitrate codecs, The Journal of the Acoustical Society of America. vol. 137. no. 65.

Li, Y., Po, L.M., Cheung, C.H., Xu, X., Feng, L., Yuan, F., & Cheung, K.W. (2016). No-Reference Video Quality Assessment With 3D Shearlet Transform and Convolutional Neural Networks, IEEE Transactions on Circuits and Systems for Video Technology. vol. 26. no. 6. pp. 1044-1057

Liu, M., Gu, K., Zhai, G. et al. (2016). Perceptual ReducedReference Visual Quality Assessment for Contrast Alteration, IEEE Trans. on Broadcasting, pp. 1-11.

Ma, J., Liu, H., & Xu, Y. (2022). Deep learning-based image quality assessment metrics: A review. IEEE Access. 10. 19724-19740.

Mamun, N., Jassim, W.A., & Zilany, M.S.A (2015). Prediction of Speech Intelligibility Using a Neurogram Orthogonal Polynomial Measure (NOPM), IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23. no. 4. pp. 760-773

Moller, S. & Raake, A. (2014). Quality of Experience: Advanced Concepts, Applications and Methods. Springer.

Picardi, C., & Mazzola, D. (2023). Objective video quality assessment: A comprehensive review of metrics and their applications. Journal of Visual Communication and Image Representation. 89. 103622.

Yan, P., & Mou, X. (2016). Video quality assessment based on correlation between spatiotemporal motion energies. Proc. SPIE.

Yu, M. et al., (2016). Binocular perception based reducedreference stereo video quality assessment method, Journal of Visual Comm. and Image Rep. pp. 246–255

Zahid, A., & Tiago H. F. (2017). Audio-Visual Multimedia Quality Assessment: A Comprehensive Survey. IEEE Access. Vol. 5. Pg 21090-21117

Zhang, H., Zhang, L., Gao, Y., & Liu, T. (2023). A comprehensive review of neural network-based image quality assessment methods. Digital Signal Processing. 136. 103866.

Zhou, W., Jiang, G., Yu, M., Shao, F., & Peng, Z. (2014). Reduced reference stereoscopic image quality assessment based on view and disparity zero-watermarks, Signal Processing: Image Communication, vol. 29. no. 1. pp. 167–176