A comprehensive survey of digital watermarking techniques

Xuping Huang[1*]

[1]Advanced Institute of Industrial Technology
[*]Corresponding author: Xuping Huang, huang-xuping@aiit.ac.jp

Abstract Nowerdays, protecting intellectual property and ensuring the authenticity of audio content is paramount. Audio digital watermarking has emerged as a crucial technology to address these concerns. This survey paper provides a comprehensive overview of audio digital watermarking techniques, serving as a valuable resource for researchers and industry professionals alike. Digital watermarking involves the hidden embedding of information, or watermarks, within audio content. These watermarks can convey data for purposes such as authentication, copyright protection, and tracking, all while maintaining the audio's perceptual quality. The success of audio watermarking depends on balancing robustness, imperceptibility, and capacity. This survey classifies audio digital watermarking techniques into three primary domains: spatial, frequency, and time-frequency. Spatial domain techniques, including LSB insertion, spread spectrum, and quantization, are known for their simplicity and versatility. Frequency domain methods, like discrete wavelet transforms (DWT) and discrete cosine transforms (DCT), leverage spectral characteristics, providing robustness and capacity. Time-frequency domain techniques, such as Short-Time Fourier Transform (STFT) and Wavelet Transforms, excel at accommodating audio signals with diverse characteristics, making them suitable for voice recognition and audio fingerprinting. Balancing robustness and security is a pivotal aspect of audio watermarking, with trade-offs often required to withstand signal processing and malicious attacks without compromising quality. This paper explores the factors influencing these attributes and discusses security enhancement methods. The evaluation of audio watermarking techniques relies on key metrics, including Signal-to-Noise Ratio (SNR), Bit Error Rate (BER), and perceptual evaluation, serving as benchmarks for assessing watermark quality. As digital audio continues to evolve, the insights provided will guide the development of more resilient, secure, and efficient audio digital watermarking techniques, meeting the increasing demand for content protection and authentication.

Keywords digital watermarking; spatial domain watermarking; content protection; discrete cosine transforms

1 Introduction

In an era where digital audio content has proliferated, the protection of intellectual property and the assurance of content integrity are of paramount concern. Audio digital watermarking has emerged as a vital tool to address these challenges. This comprehensive survey paper endeavors to provide a thorough exploration of the diverse techniques and methodologies in audio digital watermarking, aiming to serve as a valuable resource for researchers, practitioners, and stakeholders in the multimedia industry. The digital watermarking concept, foundational to this survey, revolves around the covert embedding of information, or a watermark, into audio content. This watermark may convey essential data for authentication, copyright protection, tracking, and more, without perceptibly altering the audio's quality. By enhancing the robustness, imperceptibility, and capacity of these watermarks, audio digital watermarking plays a pivotal role in safeguarding multimedia content and ensuring its legitimacy.

1.1 Significance of digital watermarking

Audio digital watermarking holds significant importance in the contemporary digital landscape for several key reasons:

Copyright Protection: In a world where audio content is easily replicated and distributed over the internet, audio digital watermarking plays a pivotal role in protecting the intellectual property rights of content creators. It enables copyright owners to embed hidden information within audio files, asserting ownership and authenticity, thus deterring unauthorized copying and distribution.

Content Authentication: Ensuring the authenticity and integrity of audio content is crucial, particularly in contexts like audio forensics, legal proceedings, and audio evidence verification. Watermarks provide a means to verify that the audio has not been tampered with, ensuring its credibility in these applications.

Content Tracking: In a rapidly evolving digital ecosystem, content tracking and monitoring are essential. Watermarks enable the tracking of audio content across different platforms and networks, facilitating content management, monitoring, and analytics. This is especially significant for music and broadcasting industries.

Anti-Piracy Measures: The entertainment industry, including music and film, grapples with piracy issues. Digital watermarks embedded in audio content can serve as a deterrent to illegal distribution, making it easier to identify the source of pirated material and take appropriate legal action.

Data Hiding for Metadata: Audio watermarking allows for the secure and invisible embedding of metadata within audio files. This metadata can include information about the author, licensing, usage rights, and other relevant details. It ensures that this information travels with the audio content, simplifying content management and licensing.

Fighting Audio Forgery: Audio digital watermarking can assist in identifying and preventing audio forgery or deepfake technology abuse. In fields like audio authentication, speech recognition, and voice biometrics, it is crucial for ensuring that the audio data is genuine and has not been manipulated.

Content Ownership Verification: Musicians, artists, and content creators can use watermarks to claim ownership of their works. It not only acts as a deterrent against unauthorized use but also simplifies royalty tracking and distribution in the music industry.

Protection Against Distribution Manipulation: Digital watermarks provide a means to protect against manipulations that may occur during audio distribution. This includes ensuring that audio content is not altered or mixed with undesirable material as it travels through various channels.

Digital Rights Management (DRM): Audio watermarking is often integrated into DRM systems, allowing content providers to control how audio content is accessed and used by consumers. It grants the flexibility to enforce usage policies and prevent unauthorized access or sharing.

Steganography: Audio digital watermarking can be employed as a form of steganography, where hidden messages or information are embedded within audio content. This has applications in covert communication and secure data transmission.

Generally speaking, audio digital watermarking is significant for protecting intellectual property, asserting content ownership, ensuring authenticity, and facilitating content management and monitoring in an ever-evolving digital landscape. It addresses the challenges posed by unauthorized copying, distribution, manipulation, and forgery of audio content while offering new opportunities for content creators and rights holders.

Received: 2023-10-29; Published online: 2024-02-29

1.2 Fundamentals of digital watermarking

Digital watermarking is a technique used to embed hidden information or a digital signal, known as a watermark, into multimedia content such as images [1], audio, or video. The key principles of digital watermarking are based on ensuring the following attributes:

Invisibility: The primary principle of digital watermarking is that the embedded watermark should be imperceptible to human senses. In other words, the addition of the watermark should not degrade the quality of the host content to the extent that it can be noticed by viewers or listeners. This invisibility is critical to maintain the aesthetic and perceptual quality of the multimedia content.

Robustness: Digital watermarks should be able to withstand various signal processing operations, distortions, and attacks, both intentional and unintentional, without being destroyed or significantly altered. This is essential to ensure that the watermark can be reliably detected even after the host content has undergone transformations like compression, filtering, cropping, or other common operations.

Security: Security in digital watermarking involves protecting the watermark from unauthorized removal, alteration, or replacement. Security mechanisms may include encryption, authentication, and tamper detection, which ensure that only authorized parties can access and manipulate the watermark.

Payload Capacity: The amount of information that a watermark can carry is referred to as its payload capacity. The capacity of a watermark varies depending on the embedding technique and the specific characteristics of the host content. Balancing the need for high payload capacity with invisibility and robustness is a critical aspect of watermarking design.

Perceptual Models: Digital watermarking often employs perceptual models that take into account the limitations of human perception. These models help in determining where and how the watermark should be embedded to minimize the likelihood of being detected while maximizing robustness and capacity. Perceptual models consider the characteristics of the human visual or auditory system and guide the watermark embedding process accordingly.

Key Management: Many watermarking systems use cryptographic keys to control access and manipulation of the watermark. Key management is crucial for maintaining the security and integrity of the watermarking process.

Authentication: Authentication ensures that the watermark's presence and integrity can be verified by authorized parties. Authentication mechanisms help in confirming the authenticity of the watermark and its associated content.

Application-Specific Considerations: The principles of digital watermarking are often applied in a context-specific manner. For example, audio watermarking may emphasize imperceptibility and robustness while image watermarking may prioritize payload capacity.

Digital watermarking is based on the principles of imperceptibility, robustness, security, payload capacity, perceptual modeling, key management, authentication, and application-specific considerations. These principles are crucial for designing effective watermarking techniques that can protect multimedia content, assert ownership, and provide authentication and tracking capabilities while maintaining the integrity and quality of the host content.

2 Conventional works on algorithms and implementations

This paper categorizes audio digital watermarking techniques based on their key characteristics and approaches, shedding light on their unique advantages and trade-offs. Three primary domains emerge: spatial, frequency, and time-frequency. The spatial domain techniques, such as LSB insertion, spread spectrum, and quantization, offer simplicity and ease of implementation, making them suitable for various applications. Frequency domain methods, which encompass discrete wavelet transforms (DWT) and discrete cosine transforms (DCT), leverage the signal's spectral properties, yielding robustness and capacity. Meanwhile, time-frequency domain techniques, like Short-Time Fourier Transform (STFT) and Wavelet Transforms, are adept at handling audio signals with varying characteristics, finding application in voice recognition and audio fingerprinting. Robustness and security are critical aspects of audio watermarking, as the content must endure various signal processing and malicious attacks while maintaining the original signal's quality. Achieving a balance between these attributes is a complex challenge, often necessitating trade-offs. We explore robustness influencing elements and security enhancing techniques, offering a detailed knowledge of the balance.

2.1 Spatial domain watermarking

Spatial domain watermarking methods involve directly manipulating the pixel values of the host image or audio signal to embed the watermark [2]. These methods are generally simple to implement but may be less robust to common image processing operations. Let's discuss some common spatial domain techniques, including LSB (Least Significant Bit) insertion, spread spectrum, and quantization, along with practical applications, advantages, and limitations.

2.1.1 LSB (Least Significant Bit) Insertion

LSB insertion is one of the most basic and widely used methods. In this approach, the least significant bit of selected pixels in the host image is replaced with the corresponding bits of the watermark data. The watermark is embedded in the least perceptible part of the image. LSB insertion is often used for copyright protection of digital images. Watermarks can be used to embed copyright information or the owner's name within the image. The LSB insertion is straightforward simple to implement. Furthermore, LSB insertion can carry a relatively large amount of watermark data, which reserves a considerable capacity for data hiding. However, as the vulnerability, LSB insertion is sensitive to common image processing operations like compression and filtering. It is fragile to attacks, such as ZeroLSB, et.al. Even minor alterations can result in the loss of the watermark.

2.1.2 Spread Spectrum

Spread spectrum watermarking involves spreading the watermark signal across the entire host signal using a pseudo-random sequence. This makes the watermark robust to various signal processing operations and attacks. This algorithm is commonly used for authentication and tamper detection in images and audio. The advantage is that the algorithm guarantees the security of the watermark by the algorithm. Furthermore, he spread spectrum technique is highly robust against common signal processing operations and attacks. The disadvantages are the complexity and the lower capacity. Implementing spread spectrum watermarking can be more complex than LSB insertion. The capacity for embedding data may be lower compared to simpler spatial domain techniques.

2.1.3 Quantization

Quantization-based watermarking embeds the watermark information by modifying the quantization levels of the host sig-

nal. This technique is widely used in audio watermarking. Quantization-based methods are suitable for embedding watermarks in audio signals for purposes like copyright protection and authentication. The advantage of this algorithm is the strong robustness. Quantization-based methods can be robust against lossy compression and various signal processing operations. The disadvantages are limited capacity and perceptibly. Quantization-based methods may have limited capacity compared to other methods, such as spread spectrum. Quantization-based watermarking can affect audio quality according to the algorithm.

As the use-case, spatial domain watermarking methods like LSB insertion, spread spectrum, and quantization offer different trade-offs in terms of simplicity, robustness, and capacity. Their practical applications vary from copyright protection to authentication and tamper detection. Choosing the appropriate method depends on the specific requirements of the application, considering factors such as robustness against common processing operations and the desired capacity for watermark data.

2.2 Frequency Domain Techniques

Frequency domain techniques in audio watermarking involve transforming the audio signal from the time domain to the frequency domain, where the watermark is embedded or extracted. This approach offers advantages such as robustness against common signal processing operations and the ability to control the perceptual impact of the watermark. Two widely used methods in the frequency domain are the Discrete Wavelet Transform (DWT) and the Discrete Cosine Transform (DCT)[3].

2.2.1 Discrete Wavelet Transform (DWT)

DWT decomposes the audio signal into various frequency components through a series of wavelet transforms. Watermark data is then embedded in selected coefficients of these components, typically in the high-frequency or detail coefficients, where changes are less perceptible. This technology can be applied for voice recognition and audio fingerprinting. DWT-based watermarking is suitable for voice recognition systems, where embedding and extracting watermarks without compromising speech intelligibility is essential. As well, audio fingerprinting, used in content identification and retrieval, can benefit from DWT-based watermarking for robustness against noise and common signal processing.

The advantages of this algorithm are the robustness and the imperceptibility. DWT can offer robustness against common signal processing operations, including compression, filtering, and noise addition. By selecting appropriate coefficients for watermark embedding, the perceptual impact on the audio quality can be minimized. The disadvantages are complexity and limited capacity. Implementing DWT-based watermarking can be computationally intensive, particularly for real-time applications. The capacity for watermark data may be limited, depending on the selected coefficients for embedding.

2.2.2 Discrete Cosine Transform (DCT)

DCT is a frequency domain transformation that converts audio signals into a representation where most of the signal's energy is concentrated in a few high or low-frequency coefficients. Watermark data is embedded in these frequency coefficients to maintain robustness. DCT-based methods [4] are suitable for embedding watermarks in audio and music files for copyright protection and authentication, and tampering detection.

The advantages of this algorithm are the robustness and efficiency. DCT-based watermarking is robust against compression [5] and common signal processing operations. DCT computations are less computationally intensive compared to some other frequency domain techniques. The disadvantages are perceptual Impact and limited capacity. Embedding watermarks in limited frequency DCT coefficients may have a more noticeable impact on audio quality compared to DWT-based methods. The capacity for watermark data is constrained by the number of different frequency DCT coefficients available for embedding.

The contextual considerations on choosing algorithms based on DWT or DCT technologies should be robustness, perceptibly, and application specific needs. The choice between DWT and DCT depends on the specific application. DWT provides more control over perceptual quality but may have limited capacity, while DCT offers better robustness but can be more perceptible. The choice of frequency domain method should align with the requirements of the audio watermarking application. For voice recognition, robustness and intelligibility are critical, favoring DWT. In music copyright protection, DCT may be preferred for its efficiency and robustness.

2.3 Time-Frequency Domain Techniques

Time-frequency domain techniques in audio watermarking are particularly suited for signals with varying characteristics and offer a way to embed and extract watermarks that are robust to time-varying signal distortions and attacks. Two prominent methods in this domain are the Short-Time Fourier Transform (STFT) and the Wavelet Transform.

2.3.1 Short-Time Fourier Transform (STFT)

The STFT is a widely used time-frequency analysis method that breaks an audio signal into small overlapping segments, then computes the Fourier transform for each segment. It results in a time-frequency representation of the signal, where both time and frequency information is preserved. STFT-based watermarking is particularly suitable for audio signals with non-stationary characteristics, where the properties of the signal change over time. These characteristics include varying pitch, amplitude, and spectral content. By segmenting the audio into smaller, time-localized components, the STFT captures these changes, making it robust against signal variations. This technology can be applied to voice recognition and audio forensics. In voice recognition systems, audio signals are often non-stationary due to changes in pitch, speed, or environmental factors. STFT-based watermarking can ensure robustness against these variations. In forensic audio analysis, where tampering or manipulation is common, STFT-based watermarking can help verify the authenticity of audio evidence.

The advantages of STFT are robustness and flexibility. STFT is resilient to time-varying distortions and attacks because it preserves the time-frequency characteristics of the audio signal. By selecting specific time-frequency components for embedding, watermarking can be tailored to balance robustness and imperceptibility. While, the disadvantages are capacity limitation and complexity. STFT-based watermarking may have limited capacity for embedding data, particularly for longer audio signals. Implementing STFT-based watermarking can be computationally intensive, especially for real-time applications.

2.3.2 Wavelet Transform

The Wavelet Transform divides an audio signal into different frequency components by using wavelet functions, each representing a different scale. This transformation provides both time and frequency information, similar to the STFT but with a different basis. The Wavelet Transform is effective in handling audio signals with varying characteristics because it decomposes the sig-

nal into various scales or resolutions, each capturing different temporal and spectral information. This makes it well-suited for non-stationary signals, where characteristics change over time. This technology can be applied to audio fingerprinting and music analysis. The Wavelet Transform is used in audio fingerprinting to create robust fingerprints that can identify audio tracks, even when subjected to alterations or distortions. Music analysis often involves dealing with complex audio signals, where characteristics like tempo, timbre, and harmony change. The Wavelet Transform helps in feature extraction and analysis.

The advantages are multi-resolution analysis and robustness. The Wavelet Transform's ability to capture signal information at multiple resolutions makes it powerful in representing audio with varying characteristics. It is robust against common signal processing operations and attacks, which is crucial for audio fingerprinting and content identification. While, the disadvantages are perceptual impact and complexity. Depending on the choice of wavelet function and decomposition parameters, the Wavelet Transform may introduce perceptible artifacts in the audio. Like the STFT, the computational complexity of the Wavelet Transform can be a limitation, particularly for real-time applications.

Possible applications for time-frequency domain based watermarking can be considered as voice recognition and audio fingerprinting. In voice recognition systems, the ability to accurately identify a speaker's voice despite variations in speech patterns, accent, pitch, and speed is essential. Time-frequency domain watermarking techniques like STFT and the Wavelet Transform are well-suited for this application. By embedding a watermark within the time-frequency components that remain consistent across different speech variations, the recognition system can verify the speaker's identity while maintaining robustness against non-stationary signal characteristics. Audio fingerprinting is used for content identification and retrieval, particularly in music and audio streaming services. The time-frequency domain techniques excel in creating robust audio fingerprints that can withstand changes in tempo, pitch, and audio quality. By embedding watermarks in these fingerprints, audio content can be tracked and identified even after alterations, such as format conversions or noise additions, have occurred. The time-frequency representations help preserve the unique characteristics of the audio, making it easier to match and identify audio content. Time-frequency domain watermarking techniques like STFT and the Wavelet Transform are well-suited for audio signals with varying characteristics. They offer robustness against time-varying distortions and attacks while finding applications in voice recognition and audio fingerprinting, where the ability to maintain signal integrity and authenticity in the presence of changing characteristics is crucial. However, the computational complexity and potential perceptual impact should be considered when choosing between these methods for specific applications. The choice between these methods depends on the specific needs of the application, considering factors like robustness, perceptibly, and computational efficiency.

3 Evaluation Literature

Evaluation metrics play a crucial role in assessing the quality and performance of audio watermarking techniques. They help quantify the effectiveness of the watermarking process, offering insights into aspects such as robustness, imperceptibility, and fidelity. Several key metrics are used to evaluate audio watermarking techniques, including Signal-to-Noise Ratio (SNR), Bit Error Rate (BER), and perceptual evaluation.

3.1 Signal-to-Noise Ratio (SNR)

SNR measures the ratio of the power of the original audio signal to the power of the noise introduced by the watermark. It is typically expressed in decibels (dB), which is an important metric in audio watermarking because it quantifies the impact of watermark embedding on the signal's quality. A high SNR indicates that the watermark has been embedded with minimal perceptual distortion, while a low SNR suggests that the watermark may be audible or affect audio quality negatively. In essence, a high SNR corresponds to better imperceptibility.

3.2 Bit Error Rate (BER)

BER calculates the discrepancy between the original watermark bits and the extracted watermark bits. It is a ratio of the number of incorrectly extracted bits to the total number of bits. It is a fundamental metric for measuring the accuracy and robustness of watermark extraction. A low BER indicates that the watermarking process is robust and reliable, as it implies that a minimal number of bits were incorrectly detected or altered during extraction. Conversely, a high BER suggests that the watermark is less robust or prone to errors during extraction.

3.3 Perceptual Evaluation

Perceptual evaluation focuses on assessing the impact of watermark embedding on the perceived quality of the audio signal. Common perceptual evaluation methods include Mean Opinion Score (MOS) and listening tests involving human listeners. Perceptual evaluation is critical in audio watermarking because it measures the perceptual quality of the watermarked audio. While SNR and BER provide objective measures, perceptual evaluation considers the human auditory system's sensitivity to audio changes. The goal is to ensure that the watermark does not introduce audible artifacts or degrade the listening experience. High-quality watermarking should maintain imperceptibility, preserving the audio's integrity.

3.4 Considerations and Trade-Offs on Evaluation

Evaluating audio watermarking techniques involves a trade-off between robustness and imperceptibility. Techniques with high SNR and low BER may compromise robustness, while maximizing robustness may lead to lower imperceptibility. The choice of metric depends on the application's specific requirements. Furthermore, perceptual evaluation metrics often include real-world testing with human listeners. These tests provide valuable insights into the subjective impact of watermarking. However, they can be time-consuming and may vary depending on the listener's preference

4 Applications of digital watermarking

This section provides practical examples of digital watermarking in details. The following application is introduced in this section for copyright protection-Online Photography Portfolio.

This is an scenario for the context of image copyright protection and ownership assertion. The achievement of the application is as follows: a professional photographer maintains an online portfolio where they showcase their high-quality images. To protect their intellectual property and assert copyright ownership, they use digital watermarking.

The implementation concerns the following phases and issues:

Watermark Embedding: The photographer selects their best images for the online portfolio and prepares high-resolution versions of these images. They then embed a digital watermark into each image using specialized software.

Watermark Content: The watermark includes information such as the photographer's name, copyright symbol, and the year of

creation. It may also include a web address to the photographer's website.

Invisibility: The photographer ensures that the watermark is added in such a way that it does not significantly detract from the viewing experience but is still clearly visible enough to identify the copyright owner.

Robustness: The watermark is designed to be robust to common manipulations that may occur during online distribution, such as resizing, cropping, and minor color adjustments.

One of the usages of this application is that When viewers or potential clients browse the portfolio, they see the watermarked images. The presence of the watermark clearly asserts the photographer's copyright ownership, making it evident that these images are protected by intellectual property laws.

By using digital watermarking technology, potential infringers are discouraged from unauthorized use or distribution of the images due to the visible watermark, as it clearly identifies the copyright owner. Ownership assertion is guaranteed since the watermark serves as a visual reminder to viewers that the images are the intellectual property of the photographer. This can help deter copyright violations and disputes. This also serves as the promotion in addition to protection, since the watermark can include the photographer's website URL, effectively promoting their brand and directing interested parties to their services. Digital watermarking technology also makes it possible to trace the abuse. In case an image is used without permission, the watermark makes it easy to trace the origin back to the photographer's portfolio. This simplifies the process of proving copyright ownership.

In this example, digital watermarking is applied to protect the intellectual property and assert copyright ownership of images in an online portfolio. It serves as a visual and traceable deterrent to unauthorized use while also promoting the photographer's brand.

However, there are several issues to consider in spite of convenience. In this case, balancing visibility and aesthetics is crucial. The watermark should be noticeable enough to assert copyright but not so obtrusive that it hinders the appreciation of the image. Furthermore, robustness is essential to ensure that the watermark remains intact even if viewers or unauthorized users manipulate the images. Moreover, the watermark should comply with copyright laws in the photographer's jurisdiction.

5 Challenges and Future Trends

5.1 Challenges and limitations

1. Robustness and Imperceptibility Trade-off
Balancing robustness against common signal processing operations and attacks with imperceptibility remains a challenge. Achieving both can be difficult, and often, watermarking techniques may need to prioritize one at the expense of the other, depending on the application.

2. High-Capacity Requirements
With the increasing demand for multimedia content protection and authentication, there is a need for high-capacity watermarking techniques that can carry substantial amounts of data. Current techniques may fall short in meeting these requirements.

3. Real-Time Processing
Some applications, such as live audio streaming or voice recognition, require real-time processing of audio signals. Implementing watermarking in real-time without introducing latency or compromising quality is a complex challenge.

4. Security and Privacy Concerns
Ensuring the security and privacy of watermarking techniques is crucial. Unauthorized access to watermarking systems, reverse engineering, and the potential misuse of watermarked content are ongoing concerns.

5. User Acceptance
In applications where audio quality is paramount, such as music and entertainment, user acceptance of watermarked content can be a challenge. Striking the right balance between imperceptibility and protection is essential.

5.2 Potential Directions for Future Research

As the development of AI and VR, the following topics might be the potential directions for future research.

1. Deep Learning and AI-Based Techniques
The application of deep learning and artificial intelligence in audio watermarking is a promising area of research. These technologies can help optimize the trade-off between robustness and imperceptibility by learning from vast amounts of data.

2. Blockchain for Copyright Management
Blockchain technology can be integrated into audio watermarking for secure and transparent copyright management. Blockchain's decentralized ledger can help establish a tamper-proof record of copyright ownership and usage rights.

3. Enhanced Real-Time Processing
Developing efficient real-time audio watermarking techniques that minimize latency and computational overhead is a critical direction for research. This can benefit applications like live streaming and voice recognition.

4. Content-Specific Watermarking
Tailoring watermarking techniques to specific content types or genres can improve both robustness and imperceptibility. For instance, watermarking for music, speech, or environmental audio can be optimized differently.

5. Cross-Modal Watermarking
Exploring the integration of watermarking techniques across different modalities, such as audio and video, can provide enhanced content protection and authentication.

6. Protecting Augmented Reality (AR) and Virtual Reality (VR)
As AR and VR technologies advance, audio watermarking will be essential to ensure the security and authenticity of spatial audio content in immersive experiences.

Digital watermarking faces challenges related to robustness, capacity, real-time processing, security, and user acceptance. Future research should explore the use of AI, blockchain, and content-specific techniques to address these challenges. Emerging trends include applications in content recognition, voice assistant security, AR/VR audio protection, smart contracts for royalties, audio forensics, and personalized content delivery. Audio watermarking is expected to play a pivotal role in securing the future of digital audio content. As the digital audio landscape continues to evolve, the insights from this survey will guide the development of more robust, secure, and efficient audio digital watermarking techniques.

References

1. M H, Lee J et al. SVD-based adaptive qim watermarking on stereo audio signals. IEEE Trans Multimedia. 2017;20: 45 – 54.

2. Wan W, Zhou K et al. JND-guided perceptually color image watermarking in spatial domain. IEEE Access. 2020;8: 164504 – 164520.

3. Byun et al S. Fast and robust watermarking method based on DCT specific location. IEEE Access. 2019;17: 100706 – 100718.

4. Ko et al H. Robust and blind image watermarking in DCT domain using inter-block coefficient correlation. Inf Sci (N Y). 2020;517: 128 – 147.

5. Wang et al. Non-aligned double JPEG compression detection based on refined markov features in QDCT domain. J Real Time Image Process. 2020;17: 7 – 16.