Department of Research & Development

Title:

SoundSafe.ai Acoustic DNA

Author: Hamad Dada

Introduction

Audio watermarking embeds imperceptible data within audio signals for purposes such as copyright protection and ownership verification. Over time, watermark removal attempts have grown in sophistication, from traditional approaches like filtering and compression to more potent threats such as AI-powered attacks. These newer attacks target the watermark itself and can be designed to evade detection by Blindfold Detection (BFD) systems.

In this paper, we explore **Acoustic DNA Reconstruction**, a self-healing watermark system designed to combat these sophisticated attacks. Inspired by biology, Acoustic DNA reconstructs damaged watermarks by leveraging fragmented data, similar to how biological DNA helps rebuild an organism's genetic code.

Watermark Embedding Considerations

To ensure the robustness of the watermark, SoundSafe.ai's psychoacoustic models embed watermarks in perceptually masked regions of the audio. This makes the watermark harder to remove without significant distortion. Spread-spectrum techniques further increase the difficulty of complete removal by distributing the watermark across different frequency or time domains. Redundancy, either through multiple variations of the same watermark or the use of multiple watermarks, further hinders complete erasure.

Watermark Self-Healing: Core Components

The core concept of watermark self-healing is to divide the watermark into **fragments**, much like how DNA is broken down into genes. Each fragment contains vital information needed to reconstruct the complete watermark.

Fragment Structure: Multi-Layered Defense

Each **Acoustic DNA** fragment contains the following components:

1. **Core Watermark Segment**: A compressed representation of the original watermark data.
2. **Error Correction Codes (ECC)**: Techniques like Reed-Solomon codes that allow for reconstruction of the watermark even if parts of it are damaged.
3. **Contextual Markers**: Metadata that indicates the position of the block within the watermark sequence, assisting in correct reconstruction order.

Comparison to Biological DNA

The **Acoustic DNA** system borrows key concepts from biological DNA. In biological DNA, genetic information is encoded in nucleotides, and redundancy within genes allows the organism to repair itself in case of mutations. Similarly, the watermark is divided into overlapping blocks, each carrying enough information to allow for reconstruction.

Acoustic DNA vs Biological DNA

Feature	Biological DNA	Acoustic DNA
Structure	Double helix of nucleotide base pairs	Linear sequence of watermark blocks
Data Representation	Nucleotides (A, T, C, G)	Core watermark, error correction, markers
Redundancy and Repair	Redundant genes repair genetic code	Overlapping blocks allow for watermark reconstruction
Reconstruction	DNA replication for cell replication	Specialized algorithms for watermark reconstruction

Mathematical Formulation of Acoustic DNA Blocks

Let WtotalW_{\text{total}}Wtotal represent the full watermark. It is composed of multiple blocks, each containing a fragment. The structure of a single block is mathematically represented as:

Wblock(t)=[Wcore(t)]⊕[Ecorr(t)]⊕[Mcontext(t)]W_{\text{block}}(t) = [W_{\text{core}}(t)] \oplus [E_{\text{corr}}(t)] \oplus [M_{\text{context}}(t)]Wblock(t)=[Wcore(t)]⊕[Ecorr(t)]⊕[Mcontext(t)]

Where:

- Wcore(t)W_{\text{core}}(t)Wcore(t) is the **Core Watermark Segment**,
- Ecorr(t)E_{\text{corr}}(t)Ecorr(t) is the **Error Correction** data,
- Mcontext(t)M_{\text{context}}(t)Mcontext(t) is the **Contextual Marker** indicating the position of the block.

Each block overlaps with adjacent blocks to ensure redundancy and resilience. The complete watermark is the sum of these blocks:

Wtotal(t)=∑i=1NWblock(i)(t)W_{\text{total}}(t) = \sum_{i=1}^{N} W_{\text{block}}^{(i)}(t)Wtotal(t)=i=1∑NWblock(i)(t)

Where NNN is the total number of blocks in the watermark.

Mathematical Breakdown of Core Components

Core Watermark Segment

The **Core Watermark Segment** is derived from the original watermark, Worig(t)W_{\text{orig}}(t)Worig(t), and is compressed to fit within the block structure. This can be mathematically represented as:

Wcore(t)=f(Worig(t),compression)W_{\text{core}}(t) = f(W_{\text{orig}}(t), \text{compression})Wcore(t)=f(Worig(t),compression)

Where:

- f(·)f(\cdot)f(·) is a compression function that reduces the size of the original watermark.

Error Correction

Reed-Solomon (RS) encoding is used for error correction to make the watermark resilient to tampering. The encoding process is represented as:

Ecorr(t)=RS(Wcore(t))E_{\text{corr}}(t) = \mathcal{RS}(W_{\text{core}}(t))Ecorr(t)=RS(Wcore(t))

Where:

- RS(·)\mathcal{RS}(\cdot)RS(·) represents the Reed-Solomon encoding that adds redundancy to the watermark.

Reed-Solomon error correction allows the recovery of the original watermark data even if a portion of the block is damaged or corrupted.

Contextual Marker

Each fragment has a **Contextual Marker**, which indicates its position in the complete watermark sequence:

Mcontext(t)=[positioni]M_{\text{context}}(t) = [\text{position}_i]Mcontext(t)=[positioni]

Where:

- positioni\text{position}_ipositioni is the index of the block within the sequence.

The **Contextual Marker** ensures the correct order of blocks during reconstruction.

Reconstruction of the Full Watermark

If part of the watermark is damaged or removed, **Acoustic DNA Reconstruction** can use the remaining fragments to reconstruct the original watermark. The reconstruction process involves utilizing the remaining blocks along with their error correction data:

Wreconstructed(t)=∑i=1Nf(Wblock(i)(t),Ecorr(i)(t))W_{\text{reconstructed}}(t) = \sum_{i=1}^{N} f(W_{\text{block}}^{(i)}(t), E_{\text{corr}}^{(i)}(t))Wreconstructed(t)=i=1∑Nf(Wblock(i)(t),Ecorr(i)(t))

Where:

- f(·)f(\cdot)f(·) represents the reconstruction function that utilizes the error correction data to recover the watermark.

In cases where some blocks are entirely missing, machine learning (ML) models can infer the missing segments based on the remaining data.

Table: Summary of Key Mathematical Components

Component	Mathematical Formula
Block Structure	Wblock(t)=[Wcore(t)]⊕[Ecorr(t)]⊕[Mcontext(t)]W_{\text{block}}(t) = [W_{\text{core}}(t)] \oplus [E_{\text{corr}}(t)] \oplus [M_{\text{context}}(t)]Wblock(t)=[Wcore(t)]⊕[Ecorr(t)]⊕[Mcontext(t)]
Core Watermark Segment	Wcore(t)=f(Worig(t),compression)W_{\text{core}}(t) = f(W_{\text{orig}}(t), \text{compression})Wcore(t)=f(Worig(t),compression)
Error Correction	Ecorr(t)=RS(Wcore(t))E_{\text{corr}}(t) = \mathcal{RS}(W_{\text{core}}(t))Ecorr(t)=RS(Wcore(t))
Contextual Marker	Mcontext(t)=[positioni]M_{\text{context}}(t) = [\text{position}_i]Mcontext(t)=[positioni]
Full Watermark	Wtotal(t)=∑i=1NWblock(i)(t)W_{\text{total}}(t) = \sum_{i=1}^{N} W_{\text{block}}^{(i)}(t)Wtotal(t)=∑i=1NWblock(i)(t)
Reconstructed Watermark	Wreconstructed(t)=∑i=1Nf(Wblock(i)(t),Ecorr(i)(t))W_{\text{reconstructed}}(t) = \sum_{i=1}^{N} f(W_{\text{block}}^{(i)}(t), E_{\text{corr}}^{(i)}(t))Wreconstructed(t)=∑i=1Nf(Wblock(i)(t),Ecorr(i)(t))