JAIST Repository
https://dspace.jaist.ac.jp/

Title	楽曲電子指紋の時間的連続性に着目した省メモリ化とデータベースの実時間検索ハードウェアアクセラレーション
Author(s)	野本，健心
Citation	
Issue Date	2023-03
Type	Thesis or Dissertation
Text version	author
URL	http://hdl.handle.net/10119/18324
Rights	
Description	Supervisor: 井口 寧，先端科学技術研究科，修士(情報科学)

JAIST
JAPAN
ADVANCED INSTITUTE OF
SCIENCE AND TECHNOLOGY

Japan Advanced Institute of Science and Technology

楽曲電子指紋の時間的連続性に着目した省メモリ化と
データベースの実時間検索ハードウェアアクセラレーション

2110132 野本 健心

近年、誰もが手軽にインターネットを利用できる環境が整備され、楽曲等のコンテンツの流通において、CD 等の媒体を通したコンテンツの流通の形態に置き換わり、ネットワークを介したデジタルコンテンツの市場規模が拡大している。さらに、ネットワークを経由した楽曲の流通が盛んになったことにより、その流通の特性についても変容している。これまで EC サイトなどを通じて行われていた、原盤権を持つレコード製作者から消費者といった一方向の流通形態に加え、それまで消費者側であったユーザー間のコンテンツのやり取りといった、双方向の流通が広く行われるようになった。これまで以上に素早く情報にアクセス可能な環境が、ソーシャルメディアや配信サービスの普及により整えられ、世相や流行など、情報の鮮度が重要なコンテンツの流通においても広く利用されている。Winny などの P2P ファイル共有ソフトや、Youtube や Spotify に代表される大規模配信プラットフォームを用いたコンテンツの流通では、コンテンツへのアクセス性や拡散速度が向上した一方、複製可能性を持つデジタル情報に対する著作権の取り締まりにおいて、コンテンツ流通管理の複雑性が増したことに加え、権利者の特定には高い速度の要求がある。

このような課題に対して、電子指紋技術を用いた著作権管理のための手法が考案されている。電子指紋は、楽曲の周波数帯域ごとのエネルギー遷移など、信号データを解析し特徴量を抽出することで、数千ビットのコンパクトな表現を用いて楽曲を識別する技術である。電子指紋を用いた著作権管理のためのシステムは、ユーザー側の負担を小さくしながら、コンテンツ利用の活性化を図るため、インターネットプロバイダのルータや、配信プラットフォームのオリジンサーバなどに実装することが考えられる。その際、指紋の生成については、信号処理においてハードウェアを用いた高速な特徴量の抽出手法が知られている。しかし、指紋を用いた楽曲検索については、品質劣化した音源から生成される指紋が、元の音源から生成された指紋と完全に同一のものとはならず、類似した指紋となるといった、電子指紋特有の性質があるため、高精度で高速な検索の実現には課題が多い。

ネットワークを介してやり取りされる楽曲は、ネットワーク帯域や転送に掛かる時間、保存領域の節約などの目的で、mp3 に代表する非可逆的な手法により圧縮されていることが多いため、システムに対して、このような楽曲が入力した時、データベース内の指紋から最近傍の指紋の検索を行う必要がある。一般に、最近傍の要素検索には「次元の呪い」と呼ばれるデータベース規模に応じた計算量の爆発的な増加の課題がある。そのため、電子指紋を用いた楽曲の検索では、類似指紋の検索により、圧縮などによって品質劣化した楽曲に対する頑健性を備えている一方、探索手法の複雑化により高速化

が容易ではない。また、著作権管理システムの対象となるデータ規模は、各定額音楽配信サービスを見ても、数十～数百万楽曲以上となっており、今後も増大する大規模なデータに対して、高速な識別が要求されることが分かる。

このような課題に対して、Staged LSH と呼ばれる LSH（局所性鋭敏型ハッシュ）を用いた、確率的でありながら高速な最近傍指紋の検索手法が提案されている。加えて、手法を FPGA にオフロードすることでハードウェア上の複数の演算器を活かした高速化が達成可能であることも分かっている。一方、既存の手法では、検索精度を一定の水準に引き上げるため、システムにおいて多数のハッシュテーブルを要求することがあった。このような場合、ハッシュテーブルが指紋データベースに対して数倍程度のメモリ空間を占めるため、組み込み機器等の小規模なメモリ空間しか持たない環境で、管理可能な楽曲指紋データベース規模に大きな制限を与えていた。

このような、楽曲から生成した電子指紋の検索手法について、指紋の時間的連続性に着目した改良を行うことで、精度、検索速度、メモリ空間効率のトレードオフにおける性能の底上げを行った。評価の結果から、精度を同等とし、省メモリ化を優先したパラメータを選択を選択したうえで、FPGA デバイス上に実装した時、400 万個の指紋データベースにおいて、既存手法と比較して、約 74.99%のハッシュテーブルの削減と、約 7.03 倍の高速化を達成した。また、Staged LSH に対して改良を加えた隣接 Staged LSH において、同等の精度およびメモリ空間効率を持つパラメータを選択した場合、約 109.79 倍の高速化を達成し、提案手法が LSH を基本とした指紋の検索手法へ広く適応可能であることを示した。以上から、指紋の時間的連続性に着目した本手法が、電子指紋検索のビットエラーに対するロバスト性を維持し、メモリ空間効率の向上ならびに、高速化において有効であることが分かった。

また、ハードウェア上に実装した手法について、並列化による高速化の考案および考察を行った。既存のハッシュテーブル間の独立性に着目した並列化手法について、グローバルメモリを主要な要素の格納場所として用いる際の課題を分析し、新たにコンテンツの分割によるデータ構造を持ったコンテンツ指向の並列化手法と、複数の検索要素の同時処理を可能とするフレーム指向の並列化手法を考案した。それぞれの並列化手法を実装し、歪みの小さな指紋入力に対して 400 万個の指紋データベース上で評価を行ったところ、4 並列のコンテンツ指向の並列化では、1 並列のものと比較して約 2.16 倍の高速化を達成した。加えて、並列にアクセス可能なメモリが制限される環境において、入力される指紋の類似度が低い場合、検索速度の面でフレーム指向並列化が有効であることを確かめた。

以上のように、同等の精度を維持したまま、メモリ空間効率の向上と高速化を行い、性能の底上げをすることで、一定時間内に検索を可能とする楽曲指紋データベース規模の拡張と、限られたメモリ空間において管理可能な楽曲指紋データベース規模の拡張を行った。