

|              |   |
|--------------|---|
| Title        | 特異スペクトル分析と心理音響モデルに基づいた音響情報ハイディングとその応用   |
| Author(s)    | KARNJANA, JESSADA   |
| Citation     |   |
| Issue Date   | 2016-09   |
| Type         | Thesis or Dissertation  |
| Text version | ETD   |
| URL          | <a href="http://hdl.handle.net/10119/13824">http://hdl.handle.net/10119/13824</a> |
| Rights       |   |
| Description  | Supervisor: 鶴木 祐史, 情報科学研究科, 博士  |

|         |   |                    |                   |
|---------|---|--------------------|-------------------|
| 氏名      | JESSADA KARNJANA  |                    |                   |
| 学位の種類   | 博士(情報科学)  |                    |                   |
| 学位記番号   | 博情第 346 号   |                    |                   |
| 学位授与年月日 | 平成 28 年 9 月 23 日  |                    |                   |
| 論文題目    | Audio Information Hiding Based on Singular-Spectrum Analysis with Psychoacoustic Model and Its Applications |                    |                   |
| 論文審査委員  | 主査  | 鵜木 祐史              | 北陸先端科学技術大学院大学 准教授 |
|         |   | 赤木 正人              | 北陸先端科学技術大学院大学 教授  |
|         |   | 党 建武               | 北陸先端科学技術大学院大学 教授  |
|         |   | 田中 宏和              | 北陸先端科学技術大学院大学 准教授 |
|         |   | Wutiwivatchai Chai | NECTEC 教授         |
|         |   | Aimanee Pakinee    | タマサート大学 SIIT 助教   |

## 論文の内容の要旨

The growth of the Internet since the last century and the spread of digital-multimedia transfer are useful and convenient for us because it has enabled us to access gigantic shared data. As a consequence, demands for applications such as broadcast monitoring, owner identification, proof of ownership, transaction tracking, tampering detection, copy control, and information carrier for audio signal have increased considerably due to technologies misuse. To answer such demands, audio information hiding has been suggested. There are five requirements for audio information hiding. (1) *Inaudibility*: a property that hidden information does not affect a perceptual quality of host signals. (2) *Robustness*: an ability to extract hidden information correctly when attacks are performed. (3) *Blindness*: a property of extracting hidden information correctly without the original signal. (4) *Confidentiality*: a property of concealing the hidden data. (5) *Capacity*: quantity of hidden information. To meet the first requirement is a real challenge because the human auditory system is very sensitive. When the first and the second are required together, the challenge is tougher because they conflict with each other. Compromising them has proved to be difficult. Actually, not just two, but all requirements conflict with each other.

The aim of this research is to explore audio information hiding that can satisfy all requirements, especially the conflict between inaudibility and robustness. A literature review of various audio-information-hiding techniques has suggested that audio watermarking based on singular value decomposition (SVD) is one of the robust techniques, and the published results are promising. Fundamentally, its robustness is due to the fact that a singular value is invariant under common signal processing. The hidden information is embedded by slightly modifying singular values. However, there are two critical problems. First, the problem about the balance between inaudibility and robustness. All SVD-based schemes treat an audio signal as a meaningless time-series and rely only on a mathematical

singular-value-manipulation. They have never taken audio features or human perception into account. Second, when we see them from the acoustic-signal-processing point of view, the physical meaning of singular values has never been addressed. Thus, it seems impossible to formulate a modification rule associating with human perception. The sole philosophy behind the published embedding rules seems to be that, notwithstanding the physical meaning of singular values is unknown, a human being cannot perceive the difference between original and watermarked sounds if the modification is done slightly.

Inspired by these facts, we propose a framework based on the singular-spectrum analysis (SSA), which is closely related to the SVD. We show that, by using SSA, singular values can have the physical meaning. Actually, they are scale factors of oscillatory components of the signal. Hence, by adopting SSA, we can exploit the advantages of SVD-based techniques, and, at the same time, SSA provides us the framework in which a modification rule can be informed. Quite contrary to the philosophy of conventional SVD-based schemes, the philosophy of this work is based on an idea that the embedding rules should be based on both the nature of the audio signal and the human audio-perceptual ability. When combining human perception model, such as the psychoacoustic models, and the strength of the SVD-based technique together, it is expected that the problem of conflicting requirements can be solved. Solving this problem is the ultimate goal.

In this work, we investigate the potentiality of SSA and formulate some basic principles that can be used to achieve the goal. Six audio-information-hiding models based on SSA are proposed. The test results show that the proposed framework achieve five subgoals. The scheme we implement can keep the advantages of the SVD-based technique and, at the same time, can reach a better performance with the help of an artificial intelligence technique. We also found the connection between singular-spectrum index and peak frequency of oscillatory components and used the finding to improve performance further. In addition, the self-synchronization for watermark detection is proposed. To demonstrate that the framework is practicable, we applied it to applications of ownership protection, information carrier, and fragile audio-watermarking.

**Keywords:** audio information hiding, singular-spectrum analysis, differential evolution, psychoacoustic model, self-synchronization

## 論文審査の結果の要旨

近年、マルチメディア情報通信技術の急激な発展とともにデジタル音コンテンツのセキュリティに対するリスク（著作権保護や改ざんなど）が高まっている。音情報ハイディングは、これらのリスクを回避するために、デジタル音コンテンツの新しい情報保護技術として注目されている。この技術の利点は、利用者に知覚されないように著作権管理情報を音響信号自体に埋め込み、それを検出することで違法コピーや違法配信を防ぐ、ある

いは追跡を可能とすることにある。最近では、著作権保護だけでなく、音響信号に補助情報を埋め込み、付加価値を高める技術にも利用されている。そのため、この技術基盤を整備するためには、次の5つの要求項目を検討する必要がある。(1) 情報埋め込みに対する知覚不可能性、(2) 埋め込み情報の検出に対するブラインド性、(3) 埋め込み情報に対する頑健性、(4) 埋め込みに関する秘匿性、(5) 高い情報埋め込み率の実現。現在までに、数多くの方法が提案されてきたが、上記の要求項目をすべて満たす技術には至っておらず、中でも知覚不可能性と頑健性の間のトレードオフは技術の実現に向けて重要な検討項目になっている。

本研究では、これらの要求項目を満たすために、データ駆動型の分析手法である特異スペクトル分析 (SSA) に基づいた音情報ハイディング手法の体系化を図った。まず、従来の特異値分解 (SVD) に基づく音情報ハイディング手法の有効性と諸問題を洗い出し、知覚不可能性と頑健性のトレードオフを解決するための原案を提案した。次に、この原案に基づいた SSA ベースの音情報ハイディング法を提案し、全要求項目を満たすための最適なパラメータ設定法を提案した。特に知覚不可能性と頑健性を同時に満たすためのアイデアとして、心理音響モデルを組み込み、マスキング特性に対応して適切な特異スペクトル操作による情報埋め込みを行うことで、5つの要求項目を満たす音情報ハイディング法を実現した。最後に、提案法の応用事例として、著作権保護、秘匿情報伝送、音声改ざん検出の実現可能性を示した。これらの成果は、一般的な音情報ハイディング法として情報通信技術に大きな寄与を与えている。いずれも応用目的に応じて優位性が異なるが、特異値の選択性により、情報埋め込み強度の操作や頑健性・脆弱性の達成度も制御できる点も大きな成果である。

以上、本論文は、音の秘匿埋め込みに関する知覚不可能性と頑健性について、ヒトの知覚特性に基づき特異スペクトル分析を利用して実現したものであり、技術の応用範囲が広く、学術的に貢献するところが大きい。よって博士 (情報科学) の学位論文として十分価値あるものと認めた。