Title	
Author(s)	Ngo, Minh Nhut
Citation	
Issue Date	2015-09
Type	Thesis or Dissertation
Text version	ETD
URL	http://hdl.handle.net/10119/12963
Rights	
Description	Supervisor: , ,

氏　　　　　　　名	NGO MINH NHUT		
学　位　の　種　類	博士(情報科学)		
学　位　記　番　号	博情第 326 号		
学 位 授 与 年 月 日	平成 27 年 9 月 24 日		
論　　文　　題　　目	Robust Audio Data Hiding Based on Dynamic Phase Manipulation and Its Applications （動的位相変調に基づいたロバスト音響データハイディングとその応用）		
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	同	教授
	党　　建武	同	教授
	KURKOSKI Brian Michael	同	准教授
	伊藤　彰則	東北大学	教授

論文の内容の要旨

Recent years have seen a rapid development of multimedia communication technologies which facilitate our life, but at the same time put security of digital audio at many risks, such as copyright infringement, malicious tampering. Audio data hiding techniques, in which special codes are embedded into actual audio content without any affection to its normal use, have been proposed as a potential solution for these issues. In general, audio data hiding methods must satisfy five requirements: (i) inaudibility—keeping watermark imperceptible to users, (ii) blindness—avoiding using double storage and communication channels, (iii) robustness—preventing intentional attacks from illegal users, (iv) high capacity—conveying a large amount of data, and (v) high reliability—precisely detecting the watermark.

Although most reported methods could partly satisfy the requirements, the trade-off among the requirements is still challenging. To keep the watermark inaudible, it is straightforward that perceptually insensitive features of audio signals should be exploited for embedding. The resistance against modifications such as lossy compression becomes weak since the hidden data could be easily destroyed without degrading the sound quality. Finding out suitable acoustic features to ensure both inaudibility and robustness simultaneously is one of the most important tasks of audio data hiding design.

The aim of this research is to propose an audio data hiding method that achieves a reasonable trade-off among the requirements and is applicable in practical problems. This research introduces a concept of dynamic phase manipulation for audio watermarking in which the human sound-perception mechanism and sophisticated embedding rule are utilized to solve the conflict among the requirements. First, the dynamic phase manipulation scheme finds out frequency components that are insensitive to human ears and resistant against signal processing operations to keep hidden data inaudible and robust. Second, an appropriate embedding rule is employed to account for blindness and high embedding capacity. Accordingly, the proposed method of audio data hiding could obtain the inaudibility and the robustness simultaneously. The proposed method is then applied to three typical applications: copy prevention, annotated audio, and information carrier over AM radio broadcast.

The phase manipulation technique is used to embed a bit of data into an audio signal by changing the phase according to two phase patterns. The amount of phase modification is an important factor which directly decides the performance of the data hiding system on the inaudibility and the robustness. The smaller amount keeps the hidden data less audible but more weak against processing and vice versa. The main goal of this research is to find out a region of frequency components suitable for embedding and corresponding amount of phase-modification for the embedding region based on the characteristics of human auditory system (HAS) and the variability of audio signals.

According to these considerations, the dynamic phase manipulation scheme is constructed as follows. Original audio is firstly analyzed to find out suitable frequency region for embedding. The phase modification of a frequency component cause distortion in a manner that is directly proportional to the magnitude of that component. Therefore, the amount of phase-modification should also be adapted to the magnitude. The amount of phase-modification is determined based on the energy of the embedding region. The modified phase spectrum and original magnitude spectrum are processed to yield a watermarked signal. In data extraction process, the same analysis steps are performed to identify the embedding frequency components and the amount of phase-modification. Watermark decoder is then performed on the phase of these components to extract embedded data. Experimental results have shown that variant amount of phase-modification improves performance on inaudibility and robustness remarkably compared with the case that a fixed amount of phase-modification is used. It suggested that the dynamic phase manipulation scheme is effective for audio data

hiding.

The proposed framework ensures the inaudibility and the robustness by exploiting the advantages human perception mechanism and the variability of audio signals. The blindness and the embedding capacity are achieved by the nature of the embedding rules. To combat against cropping and shifting attacks, the proposed framework is built with a frame synchronization scheme. The frame synchronization is performed by searching the starting point around the size of one frame. A correct starting point is detected when the confidence of extracting a bit is the highest among all the points in that frame. Confidentiality is ensured by incorporating security parameters. Watermark is encrypted with a secret key before being embedded into audio signal.

The proposed framework is evaluated with respect to inaudibility, robustness, blindness, and embedding capacity. The inaudibility is confirmed by objective difference grade and subjective listening test. The robustness is confirmed by the accuracy of extracted data against signal processing operations and attacks. Subjective test and robustness test are also carried out to confirm the effectiveness of the dynamic phase manipulation scheme. Bit rate is varied to investigate embedding capacity as well. The proposed audio data hiding method is then applied to protecting digital audio, audio entertainment, and information carrier over AM radio broadcast.

Keywords: audio data hiding, robustness, reliability, adaptive phase modulation, quantization index modulation.

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ィ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．音響電子透かしは，これらのリスクを回避するために，ディ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４つの要求項目を注意深く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）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，位相変調に着目した知覚不可能で頑健な電子音響透かし法として二つのアイディアを提案した．一つは信号の周波数成分に対して位相偏移変調による拡散が最小になるように適応的に位相変調を行うことで，知覚不可能性な形で透かし情報を埋め込む方法である．もう一つは，知覚不可能性と頑健性の両方を満たす位相スペクトルの特定領域に対して，量子指標変調を利用して透かし情報を埋め込む方法である．いずれも 4 つの要求項目を満たしており，本提案法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．特に前者の方法は，知覚不可能性と頑健性のトレードオフの観点で勝れ，後者は 102 bps という比較的高いビットレートで秘匿情報を埋め込み可能である．いずれも利用目的に応じて優位性が異なるが，将来，これらを融合して性能を更に高めることができれば，音信号に別の音を隠す秘匿通信や音信号のセキュリティを極限まで高めた通信技術に応用できる等，高い期待がもてる.

　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適応的・動的な位相変調法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.