様　式　Ｃ－１９、Ｆ－１９－１、Ｚ－１９　（共通）

科学研究費助成事業　　研究成果報告書

令和　５ 年　 ６ 月 ２３ 日現在

機関番号：２７１０１
研究種目：基盤研究(C)（一般）
研究期間：2020～2022
課題番号：２０Ｋ１１７９６

研究課題名（和文）複数の電子透かしを用いた音声による記録データの改ざん検知システムに関する研究

研究課題名（英文）A study on tampering detection system for voice signals by using multiple digital watermarks

研究代表者
上原　聡（Uehara, Satoshi）

北九州市立大学・国際環境工学部・教授

研究者番号：９０２１３３８９

交付決定額（研究期間全体）：（直接経費）　　3,200,000円

研究成果の概要（和文）：本研究課題では、録音と同時に透かし埋め込みを行うことを重視し、誤検出(改ざんが無い部分を誤って検出する場合と改ざんがある部分を検出できない場合)の割合を減らすことを目標とした。オクターブ類似性を用いた透かし埋め込みの際、一度の処理で特定周波数の信号成分を0にすることができたため録音時の透かし埋め込みが可能になった。また、エコー拡散法による透かし埋め込みをオクターブ類似性による埋め込み前に適用することで検出率が高くなることを示した。さらに計算コストとメモリへの負荷が少なく予測や解読が困難な乱数生成法を示し、その評価を行なった。乱数抽出では、タイミングを操作することで乱数列の予測をより困難にした。

研究成果の学術的意義や社会的意義
ICレコーダ等による会議内容の録音や遺言書を音声で残すことを考えると、改ざんされていない音声データであることを保証する必要がある。さらに、その音声を発した人を特定可能な音質であることも重要な要件となる。録音時の音声データに本システムを適用した透かし入り音声データは、誤差1秒以内で95%以上の改ざん箇所の特定を実現している。また、本システムによる改ざん検知では、検知対象以外の音声データが不要であることも特徴の一つとなる。

研究成果の概要（英文）：In this research project, we focused on embedding digital watermarks at the same time as recording, and we aimed to detect falsifications in the voice signal.

When embedding digital watermarks for voice signals based on octave similarity, we made it possible to zero a specific frequency component of voice signals by changing it to a negative value before IFFT. In addition, we proposed a method of generating random numbers for a watermark information without increasing computational cost and memory size, and we showed that the sequences have superior properties from results of typical randomness tests. In this system, it is more difficult to detect the watermark information for falsifications by manipulating the timing of random number extraction.

研究分野：情報セキュリティ

キーワード：音声電子透かし　改ざん検知　オクターブ類似性　エコー拡散法　整数上のロジスティック写像　ハッシュ関数　乱数抽出のタイミング

様　式　C－１９、F－１９－１、Z－１９（共通）

１．研究開始当初の背景

IC レコーダやスマートフォンの普及に伴って会話内容を記録データとして容易に録音できるようになり、会議の議事録を作成するための確認用資料としても多く用いられている。その際、重要な案件ほど音声による記録データ（音声データ）の一部または全体が編集・加工されていないことを保証する必要がある（ここで音声データの書き換えによる編集のことを「改ざん」と呼ぶ。）。しかしながら、安価な音声データ編集用ソフトウェアや無償のものでも改ざんによる編集箇所を識別するには、多くの費用と時間を費やすことになる。

２．研究の目的

本研究課題では、透かし入り音声データが改ざんされた場合に、埋め込まれた透かし情報を確認することで改ざん箇所の特定を可能とするシステムの構築を目指している。また、録音と同時に透かし情報を埋め込むことで、透かし情報の埋め込み前に音声データが編集されることを難しくしている。これより、透かし入り音声データの内容が録音時から改ざんされていないことを保証するとともに、改ざんされた場合にはその改ざん箇所を正確に特定することを目的とする。

また、本研究を遂行する上での前提条件としてもっとも重要な点（独自性）は、a. 有効な音声データであるために正当な透かし情報が埋め込まれていることであり、改ざんされたことによる透かし情報の欠如や偽りの透かし情報が改ざん箇所の特定につながることにある。さらに b. 改ざん後に正当な透かし情報を埋め込むことができない点も重要ある。また、音声を録音した後にオフラインで透かし情報の埋め込み処理を行った場合には、電子透かしの埋め込み処理を行う前に改ざんされる危険性が増えることから、c. 音声の録音と電子透かしを同時に処理するリアルタイム性が要求される。加えて透かしを埋め込んだ後も d. 話者の特定を可能とする音質保証も必要である。

３．研究の方法

有効な音声データであることを保証し、改ざんを目的とする正当な透かしの埋め込みを困難にするための条件 a ～ d を満たす A 透かし情報の生成と埋め込み処理と B 改ざん箇所の検出処理について検討する（図1.A, 1.B）。

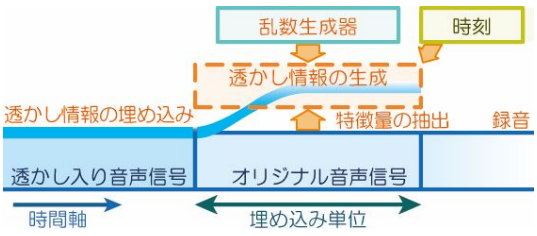

図1.A 透かし情報の生成と埋め込み処理図

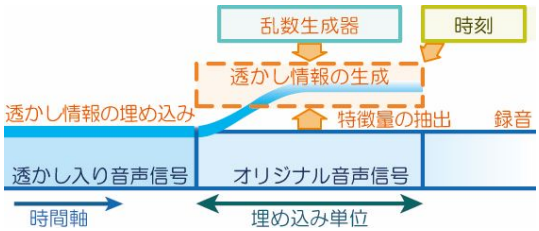

図1.B 改ざん箇所の検出処理

透かし情報の埋め込みは、サンプリング周波数44,100Hzの音声信号に対してサンプル数8192を単位として行うため，およそ0.186秒間隔で透かしの埋め込みを行うことになる。ここでは、1回の埋め込みの単位をセグメントと呼ぶ。

それぞれの処理には共通する部分もあり、5つのブロック（(1) 乱数系列の生成，(2) 音声データからの特徴量の抽出，(3) 透かし情報の生成，(4) 透かし情報の埋め込み，(5) 改ざんの検出）で構成される。次ぎに各ブロックについて整理する。

(1) 乱数系列の生成

乱数系列は透かし情報を生成する際の元データとなるため，透かし情報からの推測が困難であることが要求される。また、長時間録音にも対応するため長周期の系列であることも重要となる。さらに透かし情報の埋め込みの際にリアルタイム性も要求されるため、系列生成時の計算コストが低いことが条件となる。具体的には、解読し難い乱数性の高い系列を多数構成でき、少ないメモリで長周期の系列を高速に生成できることが要求される。本研究では、透かし情報の元になる乱数系列の生成法の提案と生成される系列の性質（安全性※，長周期性，計算コスト）についての評価を行う。※ NIST検定等による評価で問題がないことを確認する。

(2) 音声データからの特徴量の抽出

改ざんを困難にする目的から音声データの僅かな変更によって正当な透かし情報が異なる必要がある。そのため、音声データの僅かな変化でその特徴量とする値が大きく異なることが要求される。本研究では、誤検出率を低く抑えて改ざん検出率を上げることができる特徴量について考察する。目標とする具体的な両検出率はそれぞれ0%と100%であるが、研究を進める中で運用上支障の無い値についても検討したい。

(3) 透かし情報の生成

透かし情報は、3つの入力（乱数系列および音声データの特徴量、録音時刻）から生成されるが、本システムが機能するために秘密すべき情報は乱数系列であり、透かし情報から乱数系列を導出できないことが必要条件となる。また、改ざん箇所を1秒以内の誤差で検出するためには、記録データを1秒よりも小さなブロックに分割して透かし情報を埋め込む必要があり、透かし情報の生成法としてはネットワークを使用しない部分でのブロックチェーンの技法が参考になると考える。本研究では、計算コストを抑えた上で、ハッシュ関数・暗号モードなどを組み合わせた透かし情報の生成法について検討する。

(4) 透かし情報の埋め込み

透かし情報として埋め込む情報量（ビット数）が多いほど改ざんが困難になると考える。しかし、多くの情報を埋め込むことで音質が劣化し、話者の特定が困難になる点と正当な透かし情報を誤検出する原因になることも予想される。そこで本研究では、少ない情報量で改ざんを困難にする目的から、異なる電子透かしを組み合わせることを提案し、その効果について検証する。

(5) 改ざんの検出

改ざんされていない場合には正当な記録データであることを保証するために正しく透かし情報を読み取れることが重要であり、逆に改ざんされた場合には改ざん箇所を1秒以内の誤差で検出する方法について検討する。ここで、複数の電子透かしを用いる際に埋め込む順番による検出率の違いが予想される。また組み合わせる電子透かしの種類による影響もあると考える。本研究では、(4) 透かし情報の埋め込み処理と共に検討を行い、組み合わせる電子透かしの種類と透かし埋め込みの順番による検出率の差についつて考察する。また、誤り訂正の基本である多数決判定法のような手法を用いて検出率のさらなる改善についても検討する。

４．研究成果

本研究の目的である録音時に透かし情報を埋め込んで改ざんされていないことを保証することと、改ざんされた場合にはその改ざん箇所を特定可能な点は概ね達成されたと考える。特定のシステム（MATLAB）上ではあるが、録音と同時に透かし情報の埋め込みを行い、その音声データを改ざんした場合には 1 秒以内の誤差で改ざん箇所を特定できることを 30 回のシミュレーションにより 95%を超える改ざん検知率として確認している。

したがって、本研究の前提条件については、a. 有効な音声データであることを保証するための正当な透かし情報の埋め込みが可能なことと、改ざんされたことによる透かし情報の欠如や偽りの透かし情報によって改ざん箇所を特定可能なことを確認した。さらに b. 改ざん後に正当な透かし情報を埋め込むことができない点については、音声データから透かし情報を抽出した場合に録音時刻を初期値とするハッシュ関数の CTR モードを使用した系列と新しく提案した 32 ビットの乱数系列をそれぞれ抽出することが困難と考える。そこでは、音声データの特徴量に依存する乱数系列の抽出タイミングの制御を行うことで乱数列の分離を難しくしている。ただし、その詳細な評価については引き続き検討が必要である。次に c. 音声の録音と電子透かしを同時に処理するリアルタイム性については PC によるシミュレーションでは実現されているが、小型デバイスへの実装については今後実現したいと考えている。最後に d. 話者の特定を可能とする音質保証については、音声を比較することで本人と確認できるが、さらに透かしの埋め込みによる音質の劣化を抑えることが必要である。

次に研究課題を 5 つのブロックに分けて研究成果について整理する。

(1) 乱数系列の生成

小型デバイスへの実装を想定して多倍長演算を使用しない生成法を検討した。この検討課題では、カオス写像として知られるロジスティック写像を整数上で実装し、出力される0と1の出現の偏りも理解した上で、上に凸の写像と下に凸の写像を組み合わせて0と1のバランスの操作を行っている。この2つの写像を使う点では近年小型デバイス用のCPUにおいてもマルチコア化が進んでいることが背景にある。さらに、2つの写像出力値をシフト演算と排他的論理和によるビット演算で組み合わせることでCPUへの負荷軽減を図るとともに、入力値の特定を困難にする目的もある。この操作により各写像関数への入力値が32ビットに

収まらないことになるが、通常64ビット変数を使用した際に32ビットの入力値に対して上位ビットが消失することはないが、本手法では写像値が64ビットを超えた場合に下位64ビットのみ写像後の変数に残ることから逆写像が存在しないことになる。

この乱数生成法によって得られる系列は、乱数性を評価するための188項目からなるNIST検定を異なる20通りの初期値からなる乱数系列に対して、1系列当たり不合格項目の割合は0.2項目となっている。また、1Gビットの系列に対する0の出現確率は1/2から10^{-5}の誤差となっている。さらに系列の周期に対しても、他の同じ演算精度で生成される系列と比較しても十分に長い系列であり、本電子透かしに適用しても1日の録音に対して十分に長いものになっている。ただし、この生成法によって短い系列が存在することも確認できるため、今後この短い系列の生成過程についても詳しく調査する必要があると考える。最後に、この繰り返し写像による1のビットの伝搬の様子を確かめたところ、4回の写像後は出力値32ビットに広く伝搬することも確認することができた。この乱数生成法の用途は他にも多く存在すると思われるが、短い周期の存在も含めて、さらに詳細な検討が必要と考えている。

(2) 音声データからの特徴量の抽出

音声データのサンプル数8192のセグメント毎の特徴量は最大エネルギーを持つ周波数スペクトルとしている。この値は会話内容によって大きく異なるため、録音内容の改ざんによって正当な透かし情報も大きく変化することが期待できるほか、導出の際の計算コストも抑えることが可能である。この他に、第2ピークの周波数スペクトル等の使用も検討したが、誤検出による検出精度の低下の影響も大きく、最終的に最大エネルギーの周波数スペクトルのみを使用することとした。

(3) 透かし情報の生成

透かし情報は、3つの値を組み合わせて生成している。1つは録音時刻(非公開とする値でも良い)を初期値としてハッシュ関数のCTRモードを使用した256ビットの値からなる系列である。また、(1)の手法を用いた乱数性が高く計算コストが低く、演算時のメモリサイズが少ない32ビットの値からなる乱数系列を用いている。最後に、(2)の音声データの特徴量として得られる2ビットの値である。本研究では、特徴量の2ビットは透かし入りの音声データを改ざんした際に正当な透かし情報の生成を困難にするために乱数系列の抽出タイミングの制御に用いている。具体的には2ビットの特徴量により32ビットの乱数系列の出力間隔を1〜4まで変化させている。これによって元の乱数系列の特定を困難にしている。さらに、透かし情報はハッシュ値から得られる系列と乱数系列から非同期で得られる値から1部分の抽出によって得られることも、改ざんするための解析を難しくしている。

(4) 透かし情報の埋め込み

透かし情報の埋め込み方法は、埋め込み処理の負荷が少ないと考える3つの手法について検討を行なった。1つは音声信号に対する代表的な電子透かし手法であるエコー拡散法を用いている。この時、長さ4095のm-系列をカーネルに使用している。また、他の電子透かしの手法はオクターブ類似性を用いた特定周波数の振幅を操作した場合と位相を操作した場合の2種類である。本研究においてここに示す3通りの手法の組み合わせにより改ざん検知における誤検出の割合を評価した結果、エコー拡散法とオクターブ類似性を用いた特定周波数の振幅を操作する方法がもっとも誤検出が少ないことを確認できた。

(5) 改ざんの検出

各セグメントのエコー拡散法とオクターブ類似性に基づく音声電子透かしの両方が改ざんされたと判断された場合のみ、そのセグメントを改ざんが疑われるセグメントと判断する。また、改ざんが疑われるセグメントが連続して1秒以上、つまり連続する6セグメントで改ざんが疑われる場合に改ざんされたと判断している。

この検出法により、1秒以内の誤差で改ざん箇所を特定でき、改ざん検知率も95%を超えている。

５．主な発表論文等

〔雑誌論文〕　計0件

〔学会発表〕　計12件（うち招待講演　0件／うち国際学会　4件）

１．発表者名 Gen Osono, Shunsuke Araki, Takeru Miyazaki, Satoshi Uehara, Ken'ichi Kakizaki
２．発表標題 A study on control parameter and output sequence of the logistic map over integers for pseudorandom number generation
３．学会等名 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)（国際学会）
４．発表年 2021年

１．発表者名 Kako Takahashi, Takeru Miyazaki, Shunsuke Araki, Satoshi Uehara, Yasuyuki Nogami
２．発表標題 A study on digital watermarks for audio signals using phase of frequency signals in addition to the other methods
３．学会等名 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)（国際学会）
４．発表年 2021年

１．発表者名 Sota Eguchi, Takeru Miyazaki, Satoshi Uehara, Shunsuke Araki, Yasuyuki Nogam
２．発表標題 A study on relationship between period and number of divisions in piecewise logistic map over integers
３．学会等名 2021 IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW)（国際学会）
４．発表年 2021年

１．発表者名 谷口 巧実, 宮崎 武, 荒木 俊輔, 上原 聡
２．発表標題 Xorshift のパラメータと生成系列の周期の関係についての一考察
３．学会等名 第7回有限体理論とその擬似乱数系列生成への応用ワークショップ
４．発表年 2021年

１．発表者名
角井 敦志，宮崎 武，荒木 俊輔，上原 聡

２．発表標題
2つのロジスティック写像を用いた32bit演算精度の乱数生成器に関する一考察

３．学会等名
第7回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2021年

１．発表者名
宮崎 武，荒木 俊輔，上原 聡

２．発表標題
整数上のPiecewise Logistic Mapにおける分割数と系列の種類数

３．学会等名
第7回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2021年

１．発表者名
S. Eguchi, T. Miyazaki, S. Araki, S. Uehara, and Y. Nogami

２．発表標題
A study on randomness of sequences obtained from piecewise logistic map over integers

３．学会等名
2020 Intern. Symp. on Inform. Theory and Its Applications（国際学会）

４．発表年
2020年

１．発表者名
多田羅 友也，小寺 雄太，日下 卓也，野上 保之

２．発表標題
Cascaded NTU 系列の線形複雑度に関する考察

３．学会等名
第6回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2020年

１．発表者名
高谷 つぐみ，小寺 雄太，野上 保之

２．発表標題
暗号向け乱数生成における非線形フィルタの設計に関する考察

３．学会等名
第6回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2020年

１．発表者名
佐藤 陵一，小寺 雄太，野上 保之

２．発表標題
リングオシレータを用いた小規模な物理乱数生成回路の設計及び乱数性評価

３．学会等名
第6回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2020年

１．発表者名
宮崎 武，荒木 俊輔，上原 聡

２．発表標題
幾つかの系列に対する周期探索法と，素体上のロジスティック写像の逆写像演算について

３．学会等名
第6回有限体理論とその擬似乱数系列生成への応用ワークショップ

４．発表年
2020年

１．発表者名
宮崎 武、荒木 俊輔、上原 聡、野上 保之

２．発表標題
整数上のロジスティック写像による擬似乱数生成器における部分系列カオス尺度を用いた乱数性の改善法

３．学会等名
日本応用数理学会2021年研究部会連合発表会

４．発表年
2021年

〔図書〕　計0件

〔産業財産権〕

〔その他〕

―

６．研究組織

	氏名 （ローマ字氏名） （研究者番号）	所属研究機関・部局・職 （機関番号）	備考
研究分担者	野上　保之 (Nogami Yasuyuki) (60314655)	岡山大学・自然科学研究科・教授 (15301)	
研究分担者	荒木　俊輔 (Araki Shunsuke) (20332851)	九州工業大学・大学院情報工学研究院・准教授 (17104)	

７．科研費を使用して開催した国際研究集会

〔国際研究集会〕　計0件

８．本研究に関連して実施した国際共同研究の実施状況

共同研究相手国	相手方研究機関