

令和 3 年 6 月 16 日現在

機関番号：13302

研究種目：基盤研究(B)（一般）

研究期間：2017～2020

課題番号：17H01761

研究課題名（和文）メディアクローン攻撃を防御するための聴覚メディア信号処理基盤の構築

研究課題名（英文）Research on auditory-media signal processing for defending against attacks of media clones

研究代表者

鷓木 祐史（Unoki, Masashi）

北陸先端科学技術大学院大学・先端科学技術研究科・教授

研究者番号：00343187

交付決定額（研究期間全体）：（直接経費） 13,700,000円

研究成果の概要（和文）：実世界の真正データから限りなく本物に近いものとして人工的に作られたメディアは「メディアクローン」と呼ばれる。近年、このメディアクローンが実世界やサイバー空間で流通され、社会的脅威となりつつある。特に、音声合成技術を駆使して産み出された本物そっくりな音声は、「なりすまし」や「改ざん」に悪用され、音声認証システムを突破するなど重大な社会問題を巻き起こしつつある。本研究では、音声のメディアクローン攻撃に対して、適切な防御策（改ざんの検出）を実現するための聴覚メディア信号処理の基盤技術を確認した。特に、音声セキュリティに効果を持ち、メディアクローンの防御策の核となる音声情報ハイディング法を実現した。

研究成果の学術的意義や社会的意義
情報通信技術やAI技術の急激な発展により、我々を取り巻くマルチメディア情報通信の利用形態が劇的に変化し、同時に社会的脅威も急激に高まっている。特に、デジタルフォレンジクスにおける音声改ざんといった社会的脅威が高まっている。最近では、深層学習を駆使した音声合成技術の登場により、本物そっくりな合成音声を実現され、メディアクローン（ディープフェイクボイス）による事件が起こっている。音声メディアを安心安全に利用するための仕組みや、音声に含まれるプライバシー情報をメディアクローン攻撃（例えば、ディープフェイクのような攻撃）から守る革新的技術の確立が喫緊の課題である。

研究成果の概要（英文）：Non-authentic but skillfully fabricated artificial replicas of authentic media in the real world are known as “media clones.” In recent years, these media clones have become a social threat due to their increased distribution in cyber-physical spaces. In particular, media clones of voice, fabricated by using speech analysis/synthesis techniques, can be abused for spoofing and tampering with authentic speech signals. This leads to serious social problems involving, for example, fraud by voice impersonation and unauthorized speaker verification. This research project has constructed the basis of auditory-media signal processing for defending against media clone attacks. To this end, we have clarified the basic principles and fabrication of voice media clones and then propose auditory-sensing techniques consisting of auditory information hiding techniques and auditory fingerprinting techniques.

研究分野：聴覚信号処理

キーワード：聴覚メディア信号処理 メディアクローン 音声合成技術 聴覚センシング 音声情報ハイディング
音声改ざん検出 音響情報ハイディング

1. 研究開始当初の背景

現在、ネットワークを通じてやりとりされる膨大なマルチメディア情報を世界的規模で収集し、サイバーフィジカルシステムの構築やビッグデータを利活用した新たなイノベーションの創出も試みられている。これらは私たちの生活を快適で豊かなものにしてくれる反面、私たちの知らないところで個人情報を含むマルチメディア情報が悪用されるリスクも急増している。なかでも、実世界の真正データから限りなく本物に近いものとして人工的に作られたメディアは「メディアクローン」と呼ばれ、近年、このメディアクローンが実世界やサイバー空間で流通され、社会的脅威となりつつある。特に、音声合成技術を駆使して産み出された本物そっくりな音声は、「なりすまし」や「改ざん」に悪用され、音声認証システムを突破するなど重大な社会問題を巻き起こしつつある(例えば、図1)。そのため、音信号のメディアクローン攻撃に対して、適切な防御策(なりすましや改ざんの検出)を実現するための聴覚メディア信号処理の基盤技術の確立が急務である。

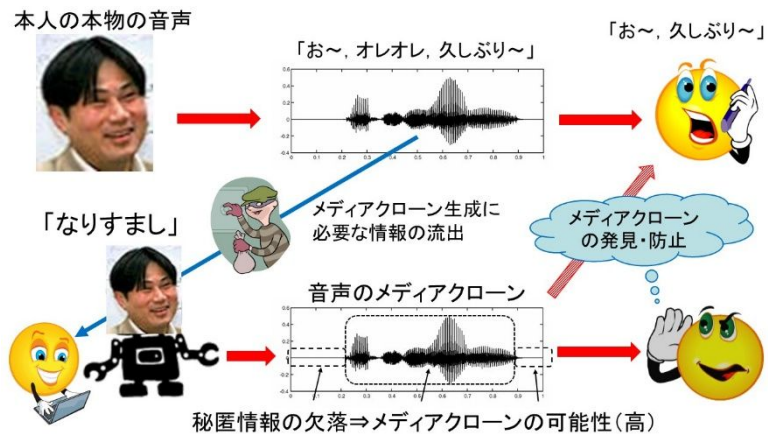


図1 音声メディアクローンの例

2. 研究の目的

本研究では、高機能な音声分析合成・編集技術の悪用を抑制し、音声のメディアクローン攻撃を防ぐための聴覚メディア信号処理の技術基盤、特に聴覚センシング技術を確立する。そのため、音声のメディアクローンの生成・認識法を深く理解した上で、聴覚的な情報ハイディング技術に音響電子指紋技術を融合させた聴覚センシングによりクローン攻撃からの防御を実現する。

3. 研究の方法

(1) 音響情報ハイディング技術

図2に一般的な音響情報ハイディング技術の概要を示す。対象となる音響信号は、埋め込み処理(符号化器)によって、付加情報を知覚不可能な形で埋め込みされる(情報秘匿される)。情報秘匿された信号は、その後、様々な経路を通り、情報機器端末等で利用者に提示される。例えば、端末がMP3プレーヤーであればMPEG(Moving Picture Experts Group)情報圧縮がかかった形で情報秘匿された信号を聴くことになるし、携帯電話やVoIP(Voice over Internet Protocol)を利用した情報通信であれば、音声符号化が施された形で情報秘匿された信号を聴くことになる。

一方、付加情報は、情報秘匿された信号から検出処理(復号器)を利用して正確に検出され、偽造防止やデータ照合に利用される。付加情報の検出に対する頑健性や脆弱性は、信号処理耐性や悪意ある攻撃への耐性など、技術利用の応用先の用途によってとらえ方がかわる。

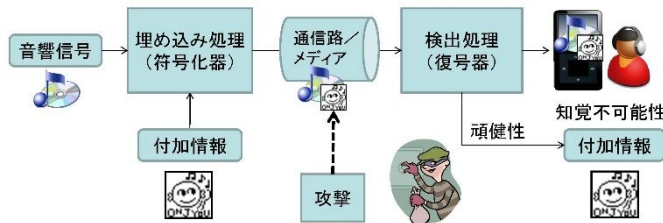


図2 音響情報ハイディング技術の概要

(2) 音声情報ハイディングへの拡張

音響情報ハイディング法は、対象を音響信号として検討しているが、音声のメディアクローンの場合、音声メディアが主対象となるため、符号化器・復号器の取り扱いについて厳密に議論されなければならない。例えば、符号化器・復号器が、MP3符号化方式であった場合、対象信号が、音楽信号はもちろん音声信号であっても問題はなく、秘匿情報の検出の頑健性そのものが常に議論される。一方、符号化器・復号器が音声符号化方式であった場合、その対象は音声に限定され、音楽信号は対象として不向きである。なぜなら、MP3符号化方式は聴覚特性を考慮して圧縮情報表現として検討されてきたのに対し、音声符号化方式は音声特有の性質と音声生成メカニズムに基づいたものであり、MP3符号化方式と異なる方向で発展してきたためである。

代表的な音声符号化方式は、符号励振線形予測(CELP: Code Excited Linear Prediction)方式である。現在、携帯電話音声に広く使われている符号化方式はAMR(Adaptive Multi-Rate)音声符号化方式である。また、インターネットなどVoIPで広く利用されている音声符号化方式は、振幅圧伸法と呼ばれる振幅圧縮方式(G.711)や適応量子化方式(G.721, G.722, G.723, G.726)などである。そのため、音声に特化した音響情報ハイディングでは、情報秘匿を行う際、対象信号そのものに埋め込むか、符号化された音声に埋め込むか、あるいは符号化器の内部で直接埋め込むなど、いくつかの方式が考えられる。

本研究では、音声メディアに特化した情報ハイディング法を音声情報ハイディング法と呼ぶ。図3に音声情報ハイディング法の分類を示す。音声符号化を経たデジタル音声データに対する情報ハイディングは、データハイディング、符号化情報ハイディング、符号化耐性情報ハイディングの三つに分類される。

データハイディング方式は、音声符号化後のデータそのものに秘匿情報を埋め込むものであり、符号化情報ハイディング方式は、音声符号化処理の中で秘匿情報を埋め込むものである。前者は音声符号化・復号の修正は不要であるが後者はそれを必要とする。また、両者とも復号された音声から秘匿情報を検出することはできない。

符号化耐性情報ハイディング方式は、従来の音響情報ハイディングと同様のものであり、音声符号化・復号処理に耐性があるものである。符号化されたもの、あるいは復号された音声のいずれからとも秘匿情報を検出できることが強みであるが、耐性強度が性能を大きく左右する。

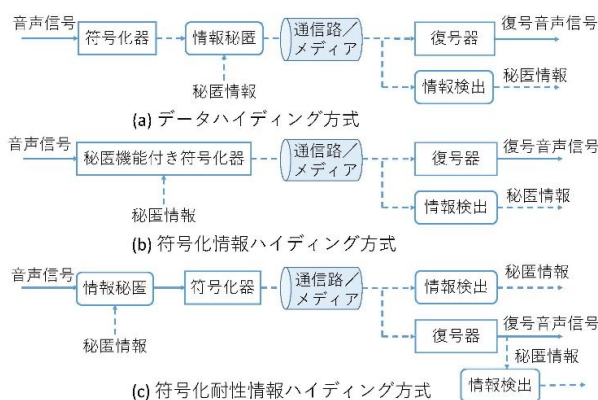


図3 三つの音声情報ハイディングの構成

データハイディング方式

データハイディング方式では、音声符号化されたデータに対して情報秘匿したことが復号後の音声品質に影響を与えることを考慮したうえで情報ハイディングを検討しなければならない。代表的な方法は、G.711 符号化音声に対する下位ビット (LSB: Least Significant Bit) 置換法や CELP 符号化や AMR でスカラ量子化されたパラメータ値に LSB 置換法を適用するものである。

音響情報ハイディングの中で最も頑健性の高いものは、スペクトル拡散に基づく方法である。この方法は、疑似乱数雑音 (PN: Pseudo-random Noise) にメッセージ情報を乗じて、対象信号に直接埋め込む方法である。秘匿情報がスペクトルに広範に埋め込まれるため、様々な処理に対して頑健性が高くなるが、スペクトル拡散による歪が音質劣化を招くという大きな問題がある。

スペクトル拡散法の埋め込み・検出のポイントは PN 信号の統計的性質 (平均が 0, 分散が 1 となる独立な確率変数) にある。本研究では、線形予測に基づく音声符号化にて、LP 残差が PN 信号に似た統計的性質を有することに着目し、線形予測に基づくスペクトル拡散型音声情報ハイディング法を提案した (図4)。この方法はノンブラインド法であるが、8 bps の情報秘匿の場合、PESQ で 4.2 程度、ビット検出率で 100%、CELP に対する耐性でビット検出率が 98% を実現した。ブラインド法についても同程度で実現可能であるが、フレーム同期が現状の課題である。

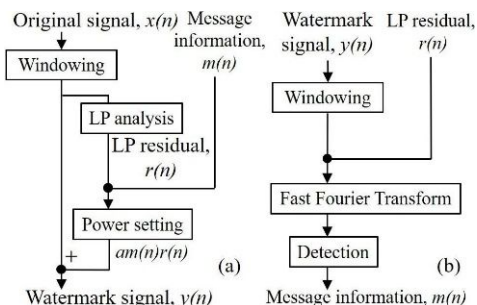


図4 スペクトル拡散型音声情報ハイディング

符号化情報ハイディング方式

音声帯域拡張用データを情報秘匿することで従来の音声符号化技術と互換性を保った高音質通話サービスに活用する案が多く提案されている。代表的なものは、CELP 系音声符号化における符号帳 (代数符号帳や固定符号帳といったコードブック) そのものに情報秘匿する方法や適応符号帳のパラメータ (ピッチ遅延データ) に情報秘匿する方法である。

線スペクトル周波数 (LSF, Line Spectral Frequency) は、CELP 系符号化と整合性がよく、耐性があることがわかっている。本研究では、図5に示すように、フォルマント操作の一つとして LSF を利用した符号化情報ハイディング方式の音声情報

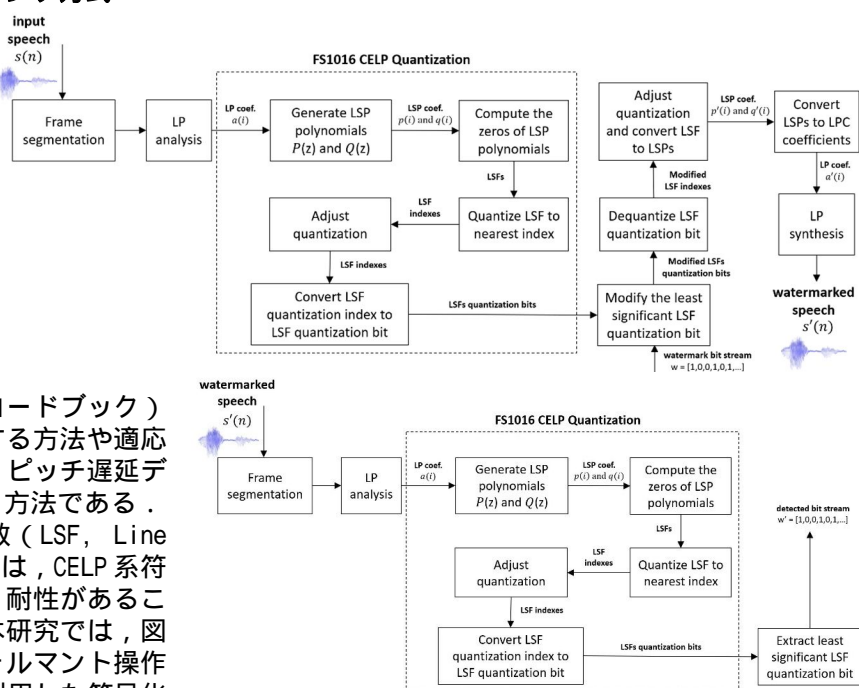


図5 LSF を利用した符号化音声情報ハイディング

ハイディングを提案した。これは、音声符号化・復号器の中で LSF の量子化指標を直接操作することで情報秘匿を狙ったものである。この方法では、8 bps の情報秘匿の場合、PESQ で 4 程度、ビット検出率が 95%、CELP に対する耐性としてビット検出率が 90% の性能を実現した。

符号化耐性情報ハイディング方式

音声符号化方式の一つは、振幅圧伸方式である。この音声符号化は、サンプルごとに瞬時に振幅を圧縮する方式 (G.711) や、過去の信号に基づく適応予測と差分信号から求められる適応量子化の両方を利用する ADPCM 方式 (G.721, G.722, G.723, G.726) に分けられる。これらに対して頑健なハイディング法は通常の音楽信号を対象としたものからでも実現できる。

本研究では、様々な観点から、音声符号化方式に頑健な、音声信号向け音響情報ハイディングを検討してきた。その中でも最良な結果をもたらすものとして、音声信号の短時間スペクトル包絡線におけるフォルマントの鋭さに情報を秘匿する方法を提案した。この方法では、図 6 に示すような考え方で、線形予測係数を LSF に変換し、LSF 対の間隔を調整することでフォルマントの鋭さを制御して音声信号に情報を秘匿する。

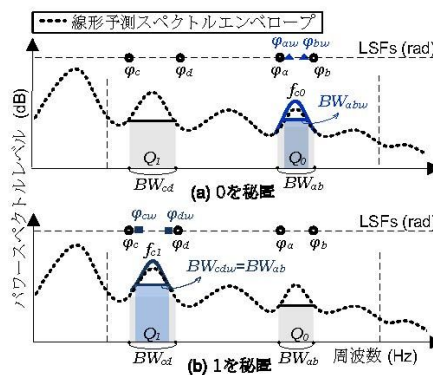


図 6 フォルマント強調処理

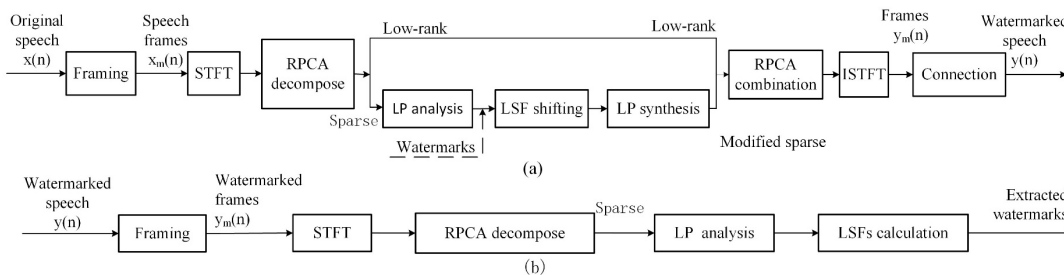


図 7 RPCA を利用したフォルマント強調型音声情報ハイディング法: (a) 埋込法と (b) 検出法

現在、この方法は、音源フィルタ理論に着目し、音源情報 (LP 残差) と声道フィルタ特性 (LP 係数に係数する LSF) にそれぞれ異なる音響情報ハイディングを適用するハイブリッド型音声情報ハイディング法に拡張されている。この方法では、音源情報に位相情報に関する QIM (Quantization Index Modulation) 法を、声道フィルタ特性にフォルマント強調に基づくハイディング法を適用することで最適な性能 (8 bps の情報秘匿の場合、PESQ で 4 程度、ビット検出率が 100%、CELP に対する耐性としてビット検出率が 85%) を発揮することが明らかにされた。

一方、フォルマント強調に基づく方法は、線形予測法が外乱に頑健ではないことに起因して、情報秘匿の耐性に課題が残されていた。ここでは、音声のスパース情報表現に基づき、雑音音声あるいは攻撃を受けた音声を音声とそれ以外に分離表現することに着目し、ロバスト主成分分析 (RPCA: Robust Principal Component Analysis) による音声表現とフォルマント強調に基づく情報ハイディング法に拡張した (図 7)。この拡張の結果、8 bps の情報秘匿の場合、PESQ で 4 程度、ビット検出率が 100%、CELP に対する耐性としてビット検出率が 95% の性能を達成した。

4. 研究成果

音声改ざん検出や音声なりすましといった重要な音声メディアのセキュリティ問題を解決するために、音声情報ハイディング技術を確認してきた。ここでは、その中でも最良な方法であるフォルマント強調を利用した改ざん検出法の性能を示す。

(1) 音声情報ハイディング法の性能

評価用音声データベースとして、ATR dataset (B セット, 日本語文章, 8.1 秒, 20 kHz サンプリング, 16bit 量子化) (Takeda et al., 2010) と LJ dataset (英語文章, 1~10 秒, 20 kHz サンプリング, 16bit 量子化) (Ito, 2017) を採用した。ATR dataset から 14 種類のサンプルを LJ dataset から 12 種類のサンプルを準備し、PESQ 並びに LSD を評価した。その結果、4~400 bps において PESQ は 4.31~3.67, LSD は 0.31~0.52 dB と提案法は知覚不可能性を十分に満たすものであった。次に埋込情報の検出におけるビット誤差率を調べたところ図 8 と図 9 のようになった。このことから提案法は十分な頑健性を有していることがわかった。

(2) 音声改ざん検出

図 10 に示すように、音声データの真正性を担保するために、音声符号化には耐性があるが、音声改ざんといった攻撃には脆弱な音声情報ハイディング法を利用する。音声改ざん検出法では、音響情報ハイディングの埋め込み部、検出部、識別部の三つの処理ブロックからなる。音声収録側では原信号と秘匿データ (真正性を担保する証拠) をペアとして保存しておき、音響情報ハイディング法を利用して、秘匿情報を原信号に埋め込み、外部へ公開するものとする。一方、ユーザ側は許可を得て流通した信号 (秘匿された信号) を取り扱うものとする。ここで、悪意のある第 3 者が公開された信号の内容そのものを、音声分析合成技術を利用して音声メディアコンテンツを改ざんし、更にそれを原信号だと偽って公開したものとする。

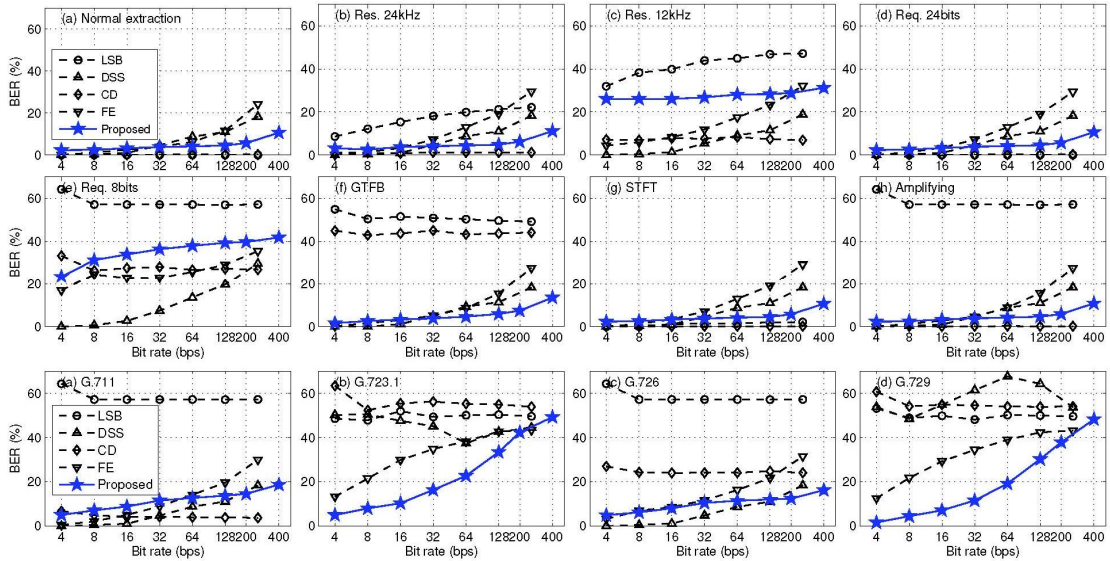


図8 RPCAを利用した音声情報ハイディング法の秘匿情報の検出（ビット誤差率，BER）

検出部では，公開された信号から埋め込み情報を検出する．識別部では収録側から秘匿情報を受け取り，検出された秘匿情報と同一であるか判断する．もし同一でない場合，第三者によって公開された信号が改ざんを受けた可能性を知ることができ，同時にその壊れ具合のパターンからどのような改ざんを受けたのかその様態を知ることができる．これらのことから，検出された透かし情報の欠落箇所からどこが改ざんされたか（改ざんの場所），どのような改ざんを受けたのか（改ざんの様態）まで検出可能である．

例えば，(1)で利用した同じ音声データについて，音声改ざんとしての9種類の攻撃（切り抜き，話速変換 50% up/down，ピッチシフト 30% up/down，高域通過/低域通過フィルタリング，残響付与 1s/3s）に対して，音声情報ハイディングの脆弱性を調べたところ，図11のようになった．図8や図9に示したように一般的な信号に対しては耐性があるものの，改ざん検出といった悪意ある処理に関しては埋込情報が検出できないことにより，改ざんの有無を判断することができる．

改ざん検出の全体性能としては，改ざんされた音声の検出に失敗した割合（False Negative Rate, %）は，9つの攻撃に対し，残響付与時に最大で6.87%，フィルタリング処理時に最小で0.0%であった．また改ざんされていない音声を改ざんされた音声と判断した割合（False Positive Rate, %）は3.65%であった．以上から，提案法により精度の高い音声改ざん検出法を実現できた．

音声改ざん検出は，音声なりすましの対策にも十分に活用できる．流通データに事前に重要な秘匿情報を音声に埋め込んでおき，それらが流通されたあとに何らかの音声処理（深層学習ベースの音声合成）によって改変された場合，それがフェイクなのかどうかを判定する（メディアクローン検出）必要がある．この課題で重要なことは，あらゆる音声符号化処理に対し頑健であり，かつなりすまし処理に関しては脆弱であるような音声情報ハイディング法を確立する必要がある．

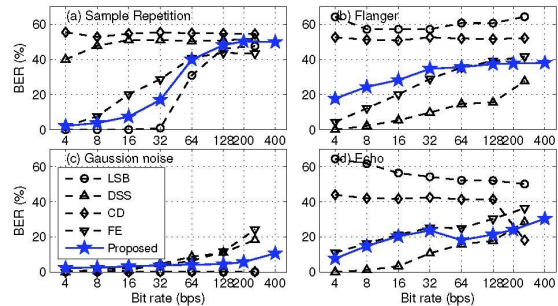


図9 信号処理に対する耐性（BER）

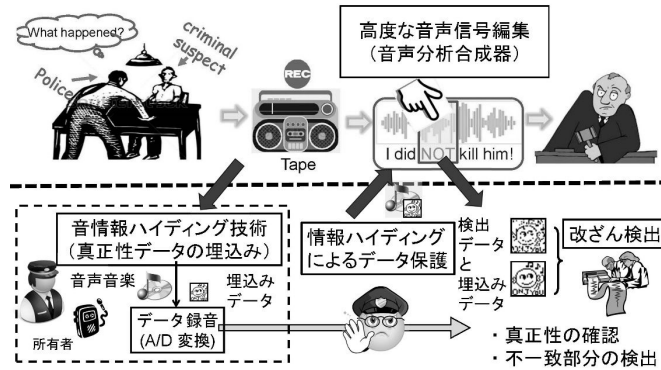


図10 攻撃に対する脆弱性（BER）

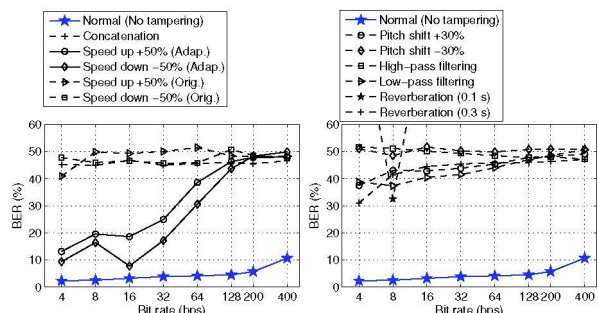


図11 音声改ざん検出におけるビット誤差率

5. 主な発表論文等

〔雑誌論文〕 計24件（うち査読付論文 21件 / うち国際共著 8件 / うちオープンアクセス 11件）

1. 著者名 Wang Shengbei, Yuan Weitao, Unoki Masashi	4. 巻 28
2. 論文標題 Multi-Subspace Echo Hiding Based on Time-Frequency Similarities of Audio Signals	5. 発行年 2020年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 2349 ~ 2363
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2020.3013785	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 Shengbei Wang, Weitao Yuan, Zhen Zhang, Jianming Wang, Masashi Unoki	4. 巻 -
2. 論文標題 Synchronous multi-bit audio watermarking based on phase shifting	5. 発行年 2021年
3. 雑誌名 Proc. ICASSP2021	6. 最初と最後の頁 2695-2699
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -
1. 著者名 Yuan Weitao, Dong Bofei, Wang Shengbei, Unoki Masashi, Wang Wenwu	4. 巻 29
2. 論文標題 Evolving Multi-Resolution Pooling CNN for Monaural Singing Voice Separation	5. 発行年 2021年
3. 雑誌名 IEEE/ACM Transactions on Audio, Speech, and Language Processing	6. 最初と最後の頁 807 ~ 822
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/TASLP.2021.3051331	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する
1. 著者名 UNOKI Masashi	4. 巻 13
2. 論文標題 Audio Information Hiding Techniques Based on Human Auditory Characteristics	5. 発行年 2020年
3. 雑誌名 IEICE ESS Fundamentals Review	6. 最初と最後の頁 284 ~ 293
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/essfr.13.4_284	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ho Tuan Vu, Akagi Masato	4. 巻 9
2. 論文標題 Cross-Lingual Voice Conversion With Controllable Speaker Individuality Using Variational Autoencoder and Star Generative Adversarial Network	5. 発行年 2021年
3. 雑誌名 IEEE Access	6. 最初と最後の頁 47503 ~ 47515
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ACCESS.2021.3063519	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 鷗木祐史	4. 巻 77(6)
2. 論文標題 音声メディアへの情報ハイディングとその応用	5. 発行年 2021年
3. 雑誌名 日本音響学会誌	6. 最初と最後の頁 382-389
掲載論文のDOI (デジタルオブジェクト識別子) 10.20697/jasj.77.6 382	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Candy Olivia Mawalim, Shengbei Wang, Masashi Unoki	4. 巻 -
2. 論文標題 Speech Information Hiding by Modification of LSF Quantization Index in CELP Codec	5. 発行年 2020年
3. 雑誌名 Proc. APSIPA2020	6. 最初と最後の頁 1321-1330
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Mawalim Candy Olivia, Unoki Masashi	4. 巻 1145
2. 論文標題 Audio Information Hiding Based on Cochlear Delay Characteristics with Optimized Segment Selection	5. 発行年 2020年
3. 雑誌名 Advances in Intelligent Systems and Computing, Springer	6. 最初と最後の頁 128-138
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-46828-6_12	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Candy Olivia Mawalim, Kasorn Galajit, Jessada Karnjana, Masashi Unoki	4. 巻 -
2. 論文標題 X-Vector Singular Value Modification and Statistical Based Decomposition with Ensemble Regression Modeling for Speaker Anonymization System	5. 発行年 2020年
3. 雑誌名 Proc. Interspeech2020	6. 最初と最後の頁 1703-1707
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Shengbei Wang, Weitao Yuan, Jianming Wang, and Masashi Unoki	4. 巻 10(14)
2. 論文標題 Speech Watermarking Based on Source-filter Model of Speech Production	5. 発行年 2019年
3. 雑誌名 Journal of Information Hiding and Multimedia Signal Processing	6. 最初と最後の頁 517-534
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Shengbei, Wang, Weitao, Yuan, Jianming, Wang, and Masashi Unoki	4. 巻 112
2. 論文標題 Detection of speech tampering using sparse representations and spectral manipulations based information hiding	5. 発行年 2019年
3. 雑誌名 Speech Communication	6. 最初と最後の頁 1~14
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.specom.2019.06.004	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Reiya Namikawa and Masashi Unoki	4. 巻 E103.D
2. 論文標題 Non-Blind Speech Watermarking Method Based on Spread-Spectrum Using Linear Prediction Residue	5. 発行年 2020年
3. 雑誌名 IEICE Transactions on Information and Systems	6. 最初と最後の頁 63~66
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/transinf.2019MUL0003	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Candy Olivia Mawalim and Masashi Unoki	4. 巻 23
2. 論文標題 Feasibility of Audio Information Hiding Using Linear Time Variant IIR Filters Based on Cochlear Delay	5. 発行年 2019年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 155 ~ 158
掲載論文のDOI (デジタルオブジェクト識別子) 10.2299/jsp.23.155	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 鷗木祐史	4. 巻 13
2. 論文標題 聴覚特性に基づいた音響情報ハイディング技術	5. 発行年 2020年
3. 雑誌名 IEICE ESS Fundamentals Review	6. 最初と最後の頁 284 ~ 293
掲載論文のDOI (デジタルオブジェクト識別子) 10.1587/essfr.13.4_284	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Boxin He, Shengbei Wang, Weitao Yuan, Jianming Wang, and Masashi Unoki	4. 巻 -
2. 論文標題 Data Augmentation for Monaural Singing Voice Separation Based on Variational Autoencoder-Generative Adversarial Network	5. 発行年 2019年
3. 雑誌名 Proc. ICME2019	6. 最初と最後の頁 1354-1359
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICME.2019.00235	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Shengbei Wang, Weitao Yuan, Jianming Wang, and Masashi Unoki	4. 巻 -
2. 論文標題 Inaudible Speech Watermarking Based on Self-compensated Echo-hiding and Sparse Subspace Clustering	5. 発行年 2019年
3. 雑誌名 Proc. ICASSP2019	6. 最初と最後の頁 2632-2636
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICASSP.2019.8682352	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Weitao Yuan, Boxin He, Shengbei Wang, Jianming Wang, Masashi Unoki	4. 巻 160
2. 論文標題 Enhanced Feature Network for Monaural Singing Voice Separation	5. 発行年 2019年
3. 雑誌名 Speech Communication	6. 最初と最後の頁 1-6
掲載論文のDOI (デジタルオブジェクト識別子) 10.1016/j.specom.2018.11.004	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 該当する

1. 著者名 Vu Ho-Tuan Vu and Masato Akagi	4. 巻 22(4)
2. 論文標題 Non-parallel Dictionary-based Voice Conversion using Variational Autoencoder with Modulation Spectrum-constrained Training	5. 発行年 2018年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 189-192
掲載論文のDOI (デジタルオブジェクト識別子) 10.2299/jsp.22.189	査読の有無 有
オープンアクセス オープンアクセスとしている (また、その予定である)	国際共著 -

1. 著者名 Kasorn Galajit, Jessada Karnjana, Aakinee Aimmanee and Masahi Unoki	4. 巻 2
2. 論文標題 Digital Audio Watermarking Method Based on Singular Spectrum Analysis with Automatic Parameter Estimation Using a Convolutional Neural Network	5. 発行年 2018年
3. 雑誌名 Recent Advances in Intelligent Information Hiding and Multimedia Signal Processing	6. 最初と最後の頁 63-73
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-030-03748-2_8	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 該当する

1. 著者名 Shengbei Wang, Weitao Yuan, Jianming Wang, Masashi Unoki	4. 巻 -
2. 論文標題 Speech watermarking based on robust principal component analysis and formant manipulations	5. 発行年 2018年
3. 雑誌名 Proc. ICASSP2018	6. 最初と最後の頁 2082-2086
掲載論文のDOI (デジタルオブジェクト識別子) 10.1109/ICASSP.2018.8462356	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Dung Kim Tran and Masashi Unoki	4. 巻 82
2. 論文標題 Study on speech representation based on spikegram for speech fingerprints	5. 発行年 2018年
3. 雑誌名 Springer Lecture Notes Smart Innovation, Systems and Technologies Volume 82: Advances in Intelligent Information Hiding and Multimedia Signal Processing	6. 最初と最後の頁 153-160
掲載論文のDOI (デジタルオブジェクト識別子) 10.1007/978-3-319-63859-1_20	査読の有無 有
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 鷗木祐史	4. 巻 317
2. 論文標題 デジタル音響信号への知覚不可能な情報ハイディング	5. 発行年 2017年
3. 雑誌名 電波技術協会報FORN	6. 最初と最後の頁 40-43
掲載論文のDOI (デジタルオブジェクト識別子) なし	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

1. 著者名 Ngo, T. V., Kubo, R., Morikawa, D., and Akagi, M	4. 巻 21(4)
2. 論文標題 Acoustical analyses of tendencies of intelligibility in Lombard speech with different background noise levels	5. 発行年 2017年
3. 雑誌名 Journal of Signal Processing	6. 最初と最後の頁 171-174
掲載論文のDOI (デジタルオブジェクト識別子) 10.2299/jsp.21.171	査読の有無 有
オープンアクセス オープンアクセスとしている(また、その予定である)	国際共著 -

1. 著者名 赤木正人	4. 巻 73(4)
2. 論文標題 コンピュータが歌うということ	5. 発行年 2017年
3. 雑誌名 日本音響学会誌	6. 最初と最後の頁 210-211
掲載論文のDOI (デジタルオブジェクト識別子) 10.20697/jasj.73.4_210	査読の有無 無
オープンアクセス オープンアクセスではない、又はオープンアクセスが困難	国際共著 -

〔学会発表〕 計32件（うち招待講演 4件 / うち国際学会 11件）

1. 発表者名 Kasorn Galajit, Jessada Karnjana, Masashi Unoki
2. 発表標題 Audio Information Hiding in Sub-signals by deploying Singular Spectrum Analysis and Psychoacoustic Model
3. 学会等名 IEICE Technical Report, EMM2020-61
4. 発表年 2021年

1. 発表者名 Candy Olivia Mawalim, Kasorn Galajit, Jessada Karnjana, Masashi Unoki
2. 発表標題 X-vector anonymization using regression modeling with statistical and singular value
3. 学会等名 IEICE Technical Report, EMM2020-62
4. 発表年 2021年

1. 発表者名 鶴木祐史, ウィン ラ ピャエ,
2. 発表標題 線形予測残差を用いたスペクトル拡散型音声情報ハイディング法の検討
3. 学会等名 IEICE Technical Report, EMM2020-63
4. 発表年 2021年

1. 発表者名 Masato Akagi
2. 発表標題 Speech communication with affective speech-to-speech translation
3. 学会等名 National Conference on Man-Machine Speech Communication (NCMMSC2019) (招待講演) (国際学会)
4. 発表年 2019年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Audio Information Hiding based on Cochlear Delay Characteristics with Optimized Segment Selection
3. 学会等名 2019 3rd International Conference on Security with Intelligent Computing and Big-data Services (SICBS 2019) (国際学会)
4. 発表年 2019年

1. 発表者名 Tuan Vu Ho and Masato Akagi
2. 発表標題 Non-parallel Voice Conversion with Controllable Speaker Individuality using Variational Autoencoder
3. 学会等名 APSIPA2019 (国際学会)
4. 発表年 2019年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Cochlear delay based audio information hiding with segment selection optimization
3. 学会等名 電子情報通信学会EMM研究会, 東北大学
4. 発表年 2019年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Study on cochlear-delay based audio information hiding by linear time-variant IIR filter
3. 学会等名 日本音響学会2019年度秋季研究発表会, 立命館大学
4. 発表年 2019年

1. 発表者名 キャンディ オリフィア マワリム, 鶴木祐史
2. 発表標題 蝸牛遅延に基づいた線形時変IIRフィルタによる音響情報ハイディング
3. 学会等名 2019年度電気・情報関係学会北陸支部連合大会, 石川高専
4. 発表年 2019年

1. 発表者名 Masashi Unoki
2. 発表標題 Audio/speech information hiding based on human auditory characteristics
3. 学会等名 10th International Conference on Knowledge and Systems Engineering (KSE 2018), Ho Chi Minh City (招待講演) (国際学会)
4. 発表年 2018年

1. 発表者名 鶴木祐史
2. 発表標題 聴覚特性に基づいた音響情報ハイディング技術
3. 学会等名 電子情報通信学会EMM研究会 (招待講演)
4. 発表年 2018年

1. 発表者名 Reiya Namikawa and Masashi Unoki
2. 発表標題 Study on speech watermarking method based on spread-spectrum using linear prediction residue
3. 学会等名 2019 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2019), (国際学会)
4. 発表年 2018年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Feasibility of Audio Information Hiding Using Linear Time Variant IIR Filter Based on Cochlear Delay
3. 学会等名 2019 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP2019), (国際学会)
4. 発表年 2018年

1. 発表者名 Kasorn Galajit, Jessada Karnjana, Aakinee Aimmanee and Masahi Unoki
2. 発表標題 Digital Audio Watermarking Method Based on Singular Spectrum Analysis with Automatic Parameter Estimation Using a Convolutional Neural Network
3. 学会等名 IHMSP2018 (国際学会)
4. 発表年 2018年

1. 発表者名 Kasorn Galajit, Mongkonchai Intarauksom, Jessada Karnjana, Pakinee Aimmaneeey, and Masashi Unoki
2. 発表標題 Speech Watermarking Technique Based on Singular Spectrum Analysis and Automatic Parameter Estimation using Differential Evolution for Tampering Detection
3. 学会等名 The 13th International Joint Symposium on Artificial Intelligence and Natural Language Processing (iSAI NLP 2018), (国際学会)
4. 発表年 2018年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Study on inaudible audio information hiding using linear-time variant IIR filter based on cochlear delay characteristics
3. 学会等名 電子情報通信学会EMM研究会, 沖縄
4. 発表年 2019年

1. 発表者名 Dung Kim Tran, Nguyen Huy Quoc, Masashi Unoki
2. 発表標題 Study on perceptual matching-pursuit algorithm to create speech representation for speech fingerprint
3. 学会等名 日本音響学会2019年度春季研究発表会講演，電気通信大学
4. 発表年 2019年

1. 発表者名 Kasorn Galajit, Jessada Karnjana, Pakinee Aimmanee, and Masashi Unoki
2. 発表標題 Study on singular spectrum analysis-based speech watermarking technique with parameter estimation using differential evolution
3. 学会等名 日本音響学会2019年度春季研究発表会講演，電気通信大学
4. 発表年 2019年

1. 発表者名 Candy Olivia Mawalim and Masashi Unoki
2. 発表標題 Study on audio information hiding using linear-time variant IIR filter based on cochlear delay characteristics
3. 学会等名 電子情報通信学会EMM研究会，東北大学RIEC
4. 発表年 2019年

1. 発表者名 Kasorn Galajit, Jessada Karnjana, Pakinee Aimmanee, and Masashi Unoki
2. 発表標題 Study on Digital Audio Watermarking Method Based on Singular Spectrum Analysis with Automatic Parameter Estimation Using a Convolutional Neural Network
3. 学会等名 電子情報通信学会EMM研究会，東北大学RIEC
4. 発表年 2019年

1. 発表者名 並河伶弥, 鷗木祐史
2. 発表標題 線形予測残差を用いたスペクトル拡散型電子透かし法の検討
3. 学会等名 電子情報通信学会EMM研究会, ホテルこうしゅうえん
4. 発表年 2018年

1. 発表者名 Dung Kim Tran, Nguyen Huy Quoc, Masashi Unoki
2. 発表標題 Study on speech representation for speech fingerprint using perceptual matching-pursuit algorithm
3. 学会等名 電子情報通信学会EMM研究会, 別府国際コンベンションセンター
4. 発表年 2018年

1. 発表者名 Shengbei Wang, Weitao Yuan, Jianming Wang, Masashi Unoki,
2. 発表標題 Study on speech watermarking based on robust principal component analysis and formant manipulations
3. 学会等名 電子情報通信学会EMM研究会, 別府国際コンベンションセンター (国際学会)
4. 発表年 2018年

1. 発表者名 Tuan Vu Ho and Masato Akagi
2. 発表標題 Non-parallel Voice Conversion using Convolutional Variational Autoencoder
3. 学会等名 ASJ 2018 Fall Meeting, Ohita Univ.
4. 発表年 2018年

1. 発表者名 並河伶弥, 鶴木祐史
2. 発表標題 線形予測符号化に基づくスペクトル拡散型音声電子透かし法の検討
3. 学会等名 第30回電気関係学会北陸支部連合大会, JAIST
4. 発表年 2018年

1. 発表者名 並河伶弥, 鶴木祐史
2. 発表標題 線形予測符号化方式におけるスペクトル拡散型音声電子透かしの検討
3. 学会等名 電子情報通信学会EMM研究会
4. 発表年 2018年

1. 発表者名 Ngo, T. V., Kubo, R., and Akagi, M.
2. 発表標題 Acoustical rules for mimicking Lombard speech uttered in a various noise level background
3. 学会等名 日本音響学会聴覚研究会
4. 発表年 2017年

1. 発表者名 Masato Akagi
2. 発表標題 Toward Affective Speech-to-Speech Translation
3. 学会等名 日本音響学会聴覚研究会 (招待講演)
4. 発表年 2017年

1. 発表者名 Ngo, T. V., Kubo, R., and Akagi, M.
2. 発表標題 Acoustical control method for increasing intelligibility based on Lombard speech uttered in background noises with various levels
3. 学会等名 日本音響学会2017年度秋季研究発表会
4. 発表年 2017年

1. 発表者名 Ngo, T. V., Kubo, R., and Akagi, M.
2. 発表標題 Speaker-independent control model for mimicking Lombard speech uttered in background noises with various levels
3. 学会等名 日本音響学会2018年度春季研究発表会
4. 発表年 2018年

1. 発表者名 Dung Kim Tran and Masashi Unoki
2. 発表標題 Study on speech representation based on spikegram for speech fingerprints
3. 学会等名 The Thirteenth International Conference on Intelligent Information Hiding and Multimedia Signal Processing (国際学会)
4. 発表年 2017年

1. 発表者名 Ho-Tuan Vu and Akagi Masato
2. 発表標題 Non-parallel training dictionary-based voice conversion with Variational Autoencoder
3. 学会等名 2018 RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP'18) (国際学会)
4. 発表年 2018年

〔図書〕 計1件

1. 著者名 鶴木祐史, 西村竜一, 伊藤彰則, 西村明, 近藤和弘, 園田光太郎	4. 発行年 2018年
2. 出版社 コロナ社	5. 総ページ数 172
3. 書名 音響情報ハイディング	

〔産業財産権〕

〔その他〕

Science Impact https://www.ingentaconnect.com/content/sil/impact/2020/00002020/00000002/art00008

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	赤木 正人 (Akagi Masato) (20242571)	北陸先端科学技術大学院大学・先端科学技術研究科・教授 (13302)	

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
---------	---------