

令和 6 年 6 月 11 日現在

機関番号：21401

研究種目：基盤研究(C)（一般）

研究期間：2019～2023

課題番号：19K12021

研究課題名（和文）3次元音場における新たな聴覚のモデル化とAR・VRのための符号化・強調への応用

研究課題名（英文）A modeling of spatial auditory masking in 3D audio field and its application to audio coding and enhancement for AR and VR

研究代表者

西口 正之（Nishiguchi, Masayuki）

秋田県立大学・システム科学技術学部・教授

研究者番号：90756636

交付決定額（研究期間全体）：（直接経費） 3,300,000円

研究成果の概要（和文）：3次元音場における聴覚の空間的マスキング効果の作用を明らかにした。空間上に存在するある音源からの音が、他の音源からの音をどのようにマスクして聞こえにくくする効果があるのか、3次元空間中の種々の音源の配置や周波数成分について調べた。その結果、水平面上に配置された2音源間のマスキングの強さに、前頭面に関して対称性があることを見出した。一例をあげると、受聴者の真正面からの音は、その近傍からの音ばかりでなく、真後ろから来る音も聞こえにくくしているということである。更に、異なる仰角方向に配置された音源間について、天頂方向からの音は、他の方向からの音よりも強く他の音をマスクする傾向があることを確認した。

研究成果の学術的意義や社会的意義

従来、音響符号化において、聴覚のマスキング効果は主に周波数軸上のそれが利用されてきた。一方、全地球から到来する音が相互にどのように影響しあうのかを明らかにすることは、今後普及が進むであろう3次元音響やイマーシブオーディオの伝送や再生に大きな役割を果たすと考える。とくにその符号化やレンダリングにおいて、伝送情報量や処理量などのリソースの優先順位をつけるうえで欠かせない情報となる。受聴者にとって重要な情報を優先して伝送、再生することで、限られた資源のなかでより高い臨場感を得ることが出来るようになると思われる。

研究成果の概要（英文）：Spatial auditory masking in the field of 3D audio has been thoroughly investigated. This involves the measurement of how a sound signal originating from a particular direction interferes with or 'masks' another sound signal emanating from a different direction. Various arrangements or frequency components in the sound signals were utilized for measurement purposes. The results show interesting discoveries. When considering two sound signals located on the horizontal plane, it was observed that the level of masking threshold exhibits symmetry with respect to the frontal plane of the subjects. In other words, a sound signal originating directly in front of a subject not only masks signals originating nearby but also those coming from directly behind the subject. Furthermore, when examining signals positioned at different elevations, it was noted that a sound signal emanating from above the subject's head tends to induce stronger masking effects compared to signals from other directions.

研究分野：音声音響信号処理

キーワード：聴覚のマスキング効果 音響符号化 3次元音響 イマーシブオーディオ Virtual Reality Augmented Reality 空間的マスキング効果

科研費による研究は、研究者の自覚と責任において実施するものです。そのため、研究の実施や研究成果の公表等については、国の要請等に基づくものではなく、その研究成果に関する見解や責任は、研究者個人に帰属します。

### 1. 研究開始当初の背景

ある音の存在が別の音を聞こえにくくする、所謂聴覚のマスクング効果については、時間軸上の継時マスクングと、周波数軸上の同時マスクングについて古くから研究されてきた。音声あるいは音響信号の符号化には主に周波数軸上の同時マスクングが利用されてきた。一方、3次元空間の様々な方向から到来する音が相互にどのように影響しあって互いをどのように聞こえにくくしてしまうのかについては、あまり調査されてきていない。方向性マスクング解除としての研究はあるものの、空間的なマスクング効果を積極的に利用しようとする立場からの研究は殆どされてこなかった。そのような状況下、MPEG-H 3D Audio など、3次元空間上の音響信号を丸ごと符号化、伝送、再生する様なイマーシブオーディオ(3次元音響)システムの為の符号化技術の標準化が進められ、放送やゲームなどへの応用展開が始まりつつあった。

### 2. 研究の目的

3次元音響信号の効率的な符号化やレンダリングに資することを目標として、3次元空間内の音が相互にどのように影響して互いを聞こえにくくしているのか、その“空間的マスクング効果”を調査して、従来は時間-周波数的な距離のみで議論されていたマスクング効果のモデルに対して、受聴者から見た音源相互の空間的な距離(方位)という新たなパラメータを加え、より一般化されたモデルを構築することを本研究の目的とする。

### 3. 研究の方法

(1) 純音あるいは帯域雑音をマスクに、純音をマスクとして、複数のマスク周波数(中心周波数)に対して、マスクの周波数を変化させてマスクング閾値の周波数特性を調べる。マスクとマスクの水平面内での方位を変えた時に、上記マスクング閾値の周波数特性がどのように変化するかを調査する。

(2) 帯域雑音をマスクに、純音をマスクとして、複数のマスク周波数(中心周波数)に対して、マスクの周波数を変化させてマスクング閾値の周波数特性を調べる。マスクとマスクの正中面内での仰角方向を変えた時に、上記マスクング閾値の周波数特性がどのように変化するかを調査する。

(3) 複数のマスク周波数について、マスクとマスクの発生時刻を制御して、時間軸上でマスクング閾値がどのように変化するか(継時マスクング)を調査する。マスクとマスクの水平面内での方位を変えた時に、このマスクング閾値がどのように変化するかを調査する。

### 4. 研究成果

研究の方法(1)について

上記(1)について、まず受聴者と音源の方位関係を図1に示すように定義する。実験結果を図2に示す。これは、マスクに中心周波数400Hzの帯域雑音を用いて、その方位を45度とした場合のマスクング閾値を3次元的にプロットしたものである。横軸がマスクの周波数、奥行方向が方位を表している。当初我々はこの閾値にはマスクの方位(=45度方向)に1つのピークだけが存在するものと予想していたが、実験結果からは、マスク方位135度方向にも閾値

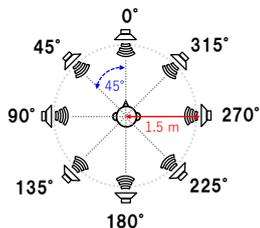


図1 音源の配置

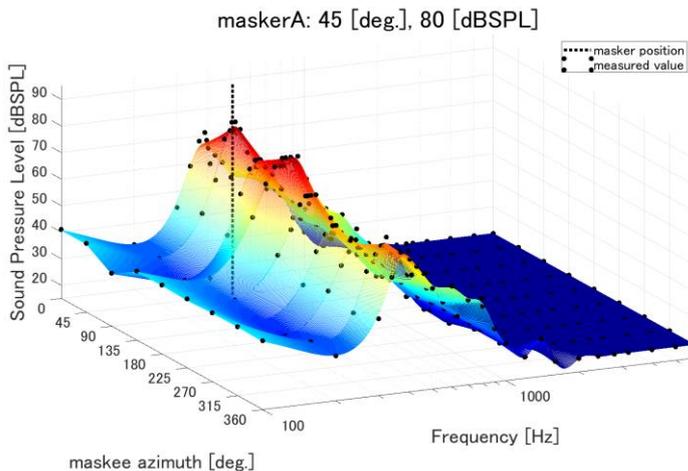


図2 マスク方向45度の時のマスクング閾値

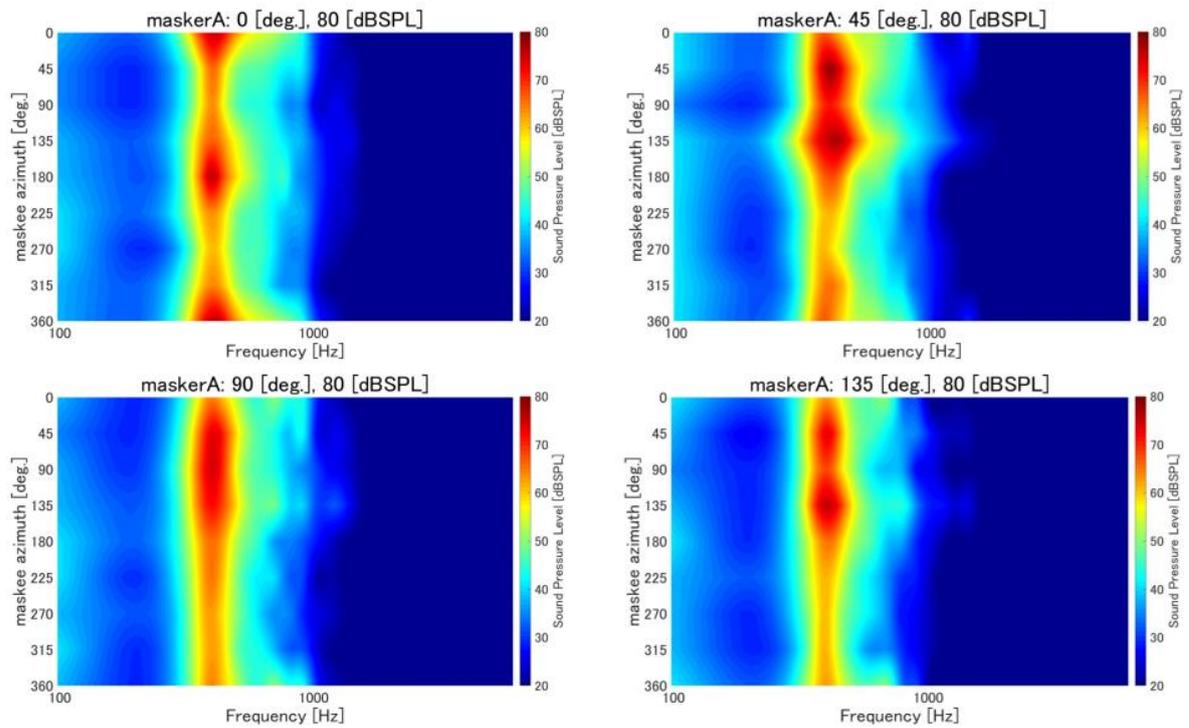


図3 各マスク方位におけるマスキング閾値

左上：マスク方位0度，右上：マスク方位45度，左下：マスク方位90度，右下：マスク方位135度

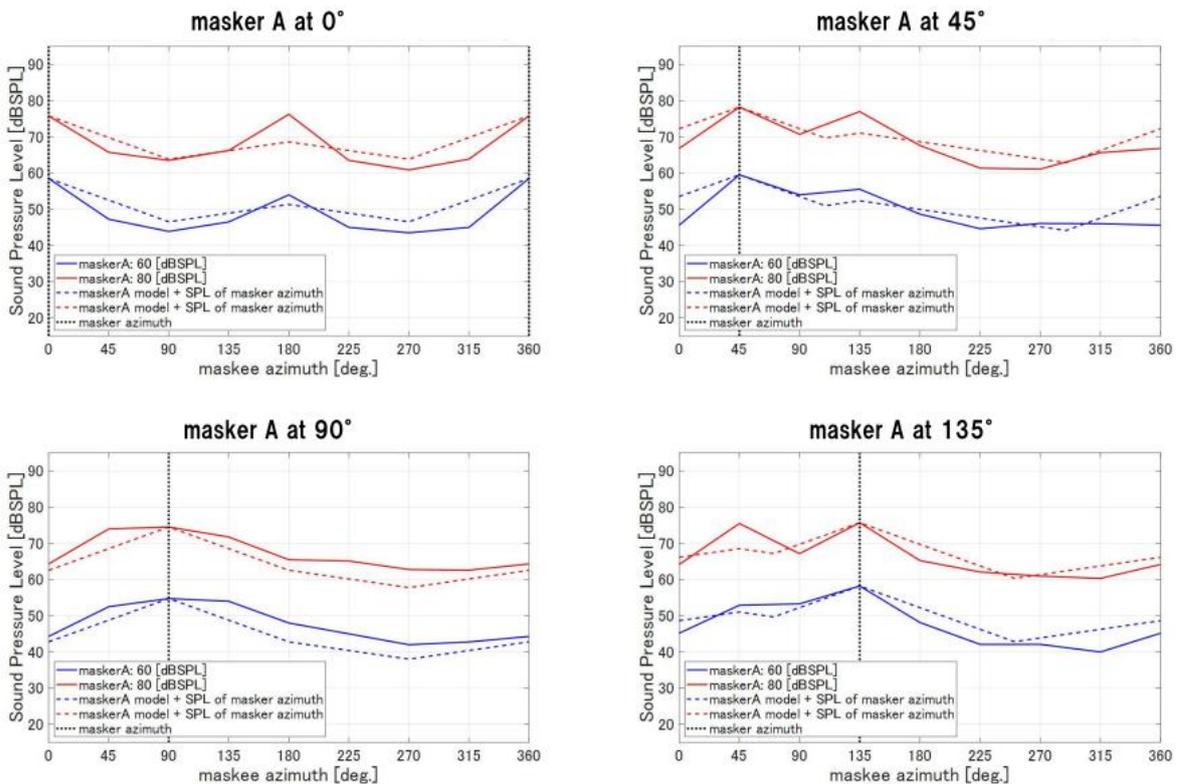


図4 各マスク方位におけるマスキング閾値のピーク

のピークが存在することがわかった。マスクの方位を0度，45度，90度，135度に変化させた時のマスキング閾値を図3に示す。図2と同様，各グラフの横軸がマスクの周波数，縦軸が方位を表している。表示されている閾値は20~80dBの範囲で，赤色は閾値の値が大きい(80dB)ことを青色は小さい(20dB)ことを示している。このグラフから，マスクが0度方向に在る場合は0度近傍に加えて180度方向の音が，マスクが45度方向に在る場合は45度近傍に加えて135度方向の音が，マスクが135度方向にある場合は135度近傍に加えて45度方向の音が強くマスクされている(=閾値が上昇している)ことがわかる。つまり，受聴者の左右の耳を結ぶ線に関して線対称の位置にもマスクが存在するかのような閾値の形状になっていることが明らかになっ

た。これを、マスキング閾値の前頭面に関する対称性と呼ぶことにする。図2の閾値の尾根、すなわち図3の各マスキ方位における周波数軸上の閾値のピーク値を拾ってプロットしたものを図4に示す。赤線がマスカレベルが80dB、青線が60dBの時の値である。実線は実測値、点線は実測値から導出したモデル式によって得た近似値である。このように、マスカ方位によって概ねマスキング閾値のピークを推定することが可能になった。これに従来からある周波数軸上のマスキング閾値のモデル式を組み合わせることで、全方位から到来する音を考慮した上で、ある方位における閾値を推定することが概ね可能になったと考える。

2chのステレオ信号にこのモデルを適用したところ、実在するあるCDの音源で、そのPerceptual Entropy(聴感上の情報量)が約5%程度削減されていることを確認した。これは符号化のビットレートを聴感上の劣化なしに5%程度削減できることを示唆している。

#### 研究の方法(2)について

(1)の実験では、水平面内に配置された音源間のマスキング効果について調査した。一方、3D音響システムを構築する際には、様々な高さ方向からの音を伝送、再生する必要がある。そこで異なる仰角に配置された2音源間のマスキング効果について調査をした。図5にマスカとマスキの位置関係を示す。図6、図7に結果を示す。図6は中心周波数400Hzの帯域雑音を用いた時の、図7は同じく1kHzの帯域雑音を用いた時の結果である。横軸が周波数、奥行方向がマスキの仰角である。ともにマスカの仰角は0度(受聴者の真正面)である。マスカの中心周波数が400Hzの時には、マスキ仰角90度近傍で尾根(各マスキ仰角における周波数軸上の閾値のピークを繋いだもの)がV字型にへこんでいるのに対して、マスカの中心周波数が1kHzのときにはそのようなへこみははなく閾値の尾根はほぼフラットである。図8にこの尾根をプロットしたものを示す。横軸はマスキの仰角、縦軸は閾値である。上のグラフはマスカ中心周波数400Hz、下のグラフは1kHzのものである。また両方のグラフともに、青線はマスカの仰角0度、赤線はマスカの仰角90度の時のものである。図8より、赤線の方が青線より上方にあることから、仰角90度(天頂)方向からのマスカは0度方向からのマスカよりもマスキの仰角に関わらず閾値が上昇する、つまり他の信号をマスクし易くなることが分かる。また、マスカの中心周波数が400Hzの時には、マスカの方位が0度90度のいずれの場合も、マスキが90度(天頂)付近で閾値が下がる、つまりマスクされにくくなることが分かる。つまり天頂方向からの信号は、他をマスクし易く自身はマスクされにくい傾向があるということが分かった。

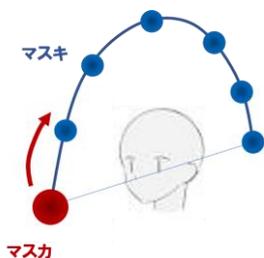


図5 異なる仰角に配置されたマスカとマスキ

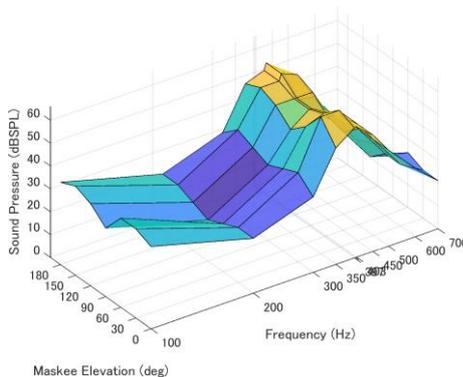


図6 マスキの仰角による閾値の変化  
マスカ仰角0度 周波数400Hz

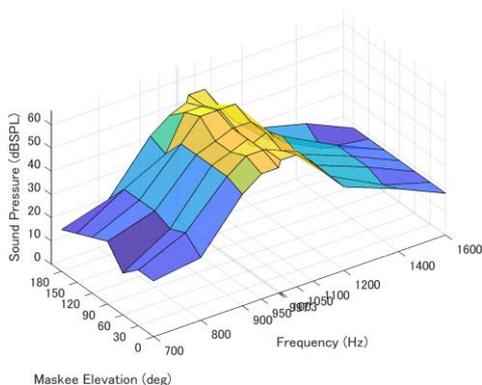


図7 マスキの仰角による閾値の変化  
マスカ仰角0度 周波数1kHz

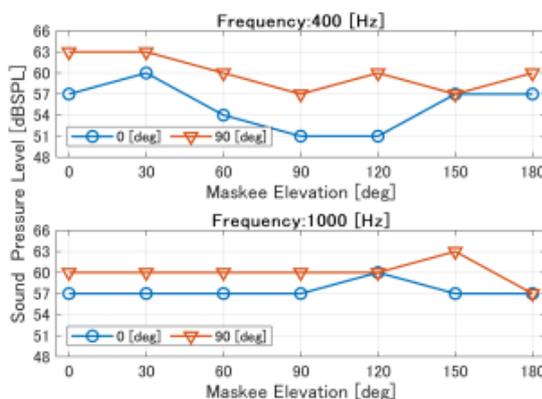


図8 マスキの仰角による閾値の変化  
青線:マスカ仰角0度 赤線:マスカ仰角90度

研究の方法(3)について

水平面上の異なる方位に配置されたマスクとマスクに関して、それらの配置によってその継時マスクングがどのように変化するか調査をした。使用した音源を表 1 に示す。頭部中心位置でのマスクの音圧レベルを 60 dB SPL とし、マスクの音圧レベルを 60~18 dB SPL の間を 3 dB 刻みで変化させた。マスクの方位を 45° に固定し、マスクは 0° ~315° までの計 8 方向について継時マスクングの閾値を測定した。刺激音のシーケンスを図 9 に示す。マスク音のみを提示することでマスクがマスクによって完全にマスクされている状態を再現した 1 つ目のシーケンスと、実際にマスクが存在する 2 つ目のシーケンスの音色を比較して評価をした。2 つ目のシーケンスが 1 つ目のシーケンスと同じ音色に聞こえた場合は、マスクがマスクによってマスクされていると判断し、違う音色に聞こえた場合はマスクされていないと判断した。マスクとマスクの時間間隔をある値に固定したときの閾値を二分探索的に調べた。時間間隔は 0, 1, 3, 5, 10, 15, 25, 50, 100, 200 [ms] の 10 条件とした。

図 10 にマスク A, マスク A, 図 11 にマスク B, マスク B を用いた時の結果をそれぞれ示す。横軸にマスクとマスクの時間間隔, 奥行き方向にマスクの方位, 高さ方向に得られた閾値を示した。これより、マスクとマスクが空間的・時間的に離れていてもマスクング効果が働くことが確認された。マスクの提示方向がマスク (45°) と同方位 (45°), 及び、前頭面对称の方位 (135°) のとき、特に閾値が上昇していることが確認された。また、中心周波数が 1kHz のマスクの方が 400Hz の場合よりも全体的にマスクング効果は早く消失してしまう事を確認した。その場合でも、継時マスクングの効果が 6dB 程度低下するのに 5~10ms はかかることが確認された。このことから、空間的マスクング効果を利用した符号化において、符号化する複数の音源の空間的な位置関係が変化した場合に、6dB 程度まで聴感上の SNR 劣化を許容するのであれば、概ね 5~10ms 以内に量子化のビット数を更新すればよいと考えらる。

表 1 使用したマスク・マスク

名前	音源
マスク A	帯域雑音(中心周波数 400 Hz, 帯域幅 100 Hz)
マスク B	帯域雑音(中心周波数 1000 Hz, 帯域幅 150 Hz)
マスク A	純音(周波数 400 Hz)
マスク B	純音(周波数 1000 Hz)

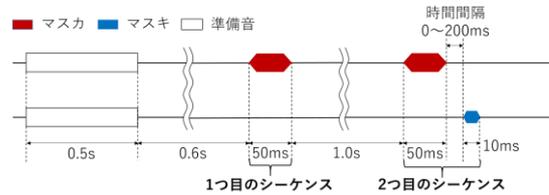


図 9 刺激音のシーケンス

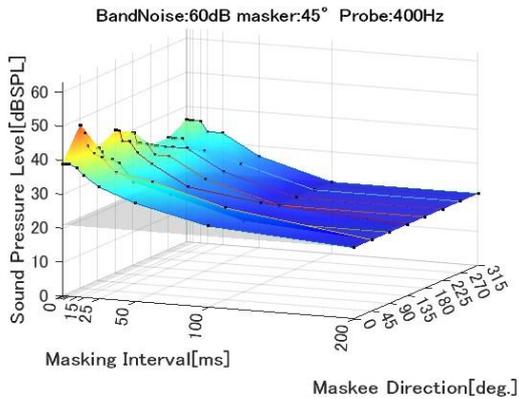


図 10 マスクの中心周波数 400Hz の時の継時マスクング

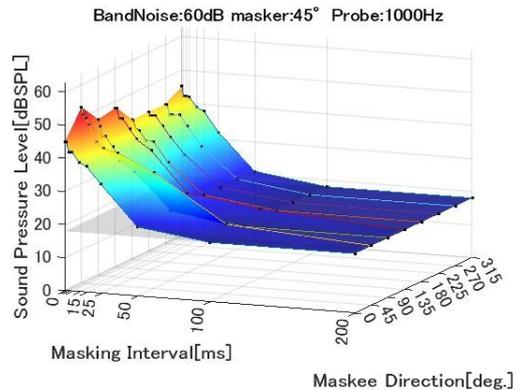


図 11 マスクの中心周波数 1kHz の時の継時マスクング

以上本研究の成果をまとめると、主に水平面内における空間的マスクング効果の作用について調査し、その閾値には前頭面に関する対称性があることを見出し、そのモデル化を行い、符号化器への応用の可能性について示した。また符号器に应用する際に、ビット割り振りの更新に必要な応答速度がおおよそ 5~10ms であることを示した。更に、異なる仰角方向の音源間のマスクング効果についてもそのふるまいについて調査を行い、天頂方向からの音が他の方向からの音よりも強く他の音をマスクすることを確かめた。

以上

5. 主な発表論文等

〔雑誌論文〕 計1件（うち査読付論文 0件 / うち国際共著 0件 / うちオープンアクセス 0件）

1. 著者名 西口正之	4. 巻 78 巻 3 号
2. 論文標題 実音源と仮想音源の間の空間的マスキングについて	5. 発行年 2022年
3. 雑誌名 日本音響学会誌	6. 最初と最後の頁 143-149
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 -

〔学会発表〕 計19件（うち招待講演 2件 / うち国際学会 6件）

1. 発表者名 Hiroyuki Fujishiro, Masayuki Nishiguchi, Kanji Watanabe, Koji Abe
2. 発表標題 Spatial auditory masking between source signals at different elevations on the median plane
3. 学会等名 155th AES convention (国際学会)
4. 発表年 2023年

1. 発表者名 Masayuki Nishiguchi, Yuuki Saito, Kanji Watanabe, Koji Abe
2. 発表標題 Spatial resolution of human hearing with different azimuths, elevations, and bandwidths of source signals
3. 学会等名 155th AES convention (国際学会)
4. 発表年 2023年

1. 発表者名 藤城博人, 西口正之, 渡邊貴治, 安倍幸治
2. 発表標題 正中面におよび矢状面上の仰角の異なる音源間の空間的マスキング効果について
3. 学会等名 応用音響研究会, 電子情報通信学会
4. 発表年 2024年

1. 発表者名 齋藤優季, 西口正之, 渡邊貫治, 安倍幸治
2. 発表標題 動的な音源の方位及び帯域が再生音の遅延の検知限に与える影響
3. 学会等名 応用音響研究会, 電子情報通信学会
4. 発表年 2024年

1. 発表者名 藤城博人, 西口正之, 渡邊貫治, 安倍幸治
2. 発表標題 正中面における仰角の異なる音源間の空間的マスクング効果
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2023年

1. 発表者名 齋藤優季, 西口正之, 渡邊貫治, 安倍幸治
2. 発表標題 音源の仰角及び帯域が音源定位分解能に与える影響
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2023年

1. 発表者名 藤城博人, 西口正之, 渡邊貫治, 安倍幸治
2. 発表標題 正中面内の異なる仰角に配置された2音源間の空間的マスクング効果について
3. 学会等名 第137回音楽情報科学研究会, 情報処理学会
4. 発表年 2023年

1. 発表者名 西口正之
2. 発表標題 聴覚の空間的マスクング効果とその応用
3. 学会等名 第40回日本ロボット学会学術講演会（招待講演）
4. 発表年 2022年

1. 発表者名 藤城博人, 西口正之, 渡邊貫治, 安倍幸治, 高根昭一
2. 発表標題 聴取者から音源までの距離の変化がマスクング効果に及ぼす影響について
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2022年

1. 発表者名 齋藤優季, 西口正之, 渡邊貫治, 安倍幸治, 高根昭一
2. 発表標題 静的・動的な音源に対する聴覚の空間的マスクング効果の継時的な特性
3. 学会等名 日本音響学会秋季研究発表会
4. 発表年 2022年

1. 発表者名 藤城博人, 西口正之, 渡邊貫治, 安倍幸治, 高根昭一
2. 発表標題 プローブ音源を聴取者に近づけた際のマスクング効果の変化について
3. 学会等名 電気関係学会東北支部連合大会
4. 発表年 2022年

1. 発表者名 齋藤優季, 西口正之, 渡邊貫治, 安倍幸治, 高根昭一
2. 発表標題 聴覚の空間的マスクング効果の時間的振る舞いに関する検討
3. 学会等名 電気関係学会東北支部連合大会
4. 発表年 2022年

1. 発表者名 Masayuki Nishiguchi, Soma Ishihara, Kanji Watanabe, Koji Abe, and Shouichi Takane
2. 発表標題 Spatial auditory masking between real sound signals and virtual sound images
3. 学会等名 Audio Engineering Society 151th convention (国際学会)
4. 発表年 2021年

1. 発表者名 西口正之, 石原颯馬, 渡邊貫治, 安倍幸治, 高根昭一
2. 発表標題 実音源による仮想音源の空間的マスクング効果について
3. 学会等名 日本音響学会2021年秋季研究発表会
4. 発表年 2021年

1. 発表者名 Masayuki Nishiguchi, Soma Ishihara, Kanji Watanabe, koji Abe, Shouichi Takane
2. 発表標題 Spatial auditory masking of human hearing for applications of augmented reality
3. 学会等名 Asia Pacific Society for Computing and Information Technology APSCIT (招待講演) (国際学会)
4. 発表年 2021年

1. 発表者名 Masayuki Nishiguchi , Kazuki Ishimoto, Kanji Watanabe , Koji Abe, and Shouichi Takane
2. 発表標題 Spatial auditory masking caused by phantom sound images
3. 学会等名 Audio Engineering Society 149th Convention ( 国際学会 )
4. 発表年 2020年

1. 発表者名 山中佑馬, 西口正之, 渡邊貫治, 高根昭一, 安倍幸治
2. 発表標題 量子化雑音の知覚に対する空間的マスキング効果の影響
3. 学会等名 電子情報通信学会技術報告
4. 発表年 2020年

1. 発表者名 Masayuki Nishiguchi , Kodai Kato, Kanji Watanabe, Koji Abe, and Shouichi Takane
2. 発表標題 Spatial auditory masking for three-dimensional audio coding
3. 学会等名 Audio Engineering Society 147th Convention ( 国際学会 )
4. 発表年 2019年

1. 発表者名 山中佑馬 , 西口正之, 渡邊貫治, 高根昭一, 安倍幸治
2. 発表標題 空間的マスキング効果に基づく3D音響符号化に関する検討
3. 学会等名 日本音響学会 2020春季研究発表会
4. 発表年 2020年

〔図書〕 計0件

〔出願〕 計6件

産業財産権の名称 信号伝送方法、信号生成方法、信号再生方法、音声信号処理プログラム、音声伝送装置、音声再生装置、及び音声伝送再生システム	発明者 西口、齋藤	権利者 同左
産業財産権の種類、番号 特許、2023-168575	出願年 2023年	国内・外国の別 国内

産業財産権の名称 音響信号符号化方法、音響信号復号化方法、プログラム、符号化装置、音響システム、及び復号化装置	発明者 西口正之、加藤巧大	権利者 秋田県立大学
産業財産権の種類、番号 特許、特願2021-502010	出願年 2021年	国内・外国の別 国内

産業財産権の名称 音頻信号編碼方法、音頻信号解碼方法、程序、編碼装置、音頻系統及解碼装置	発明者 西口正之、加藤巧大	権利者 秋田県立大学
産業財産権の種類、番号 特許、CN 113574596A	出願年 2021年	国内・外国の別 外国

産業財産権の名称 Acoustic signal encoding method, acoustic signal decoding method, program, encoding device, acoustic system, and decoding device	発明者 西口正之、加藤巧大	権利者 秋田県立大学
産業財産権の種類、番号 特許、EP20759801.2	出願年 2021年	国内・外国の別 外国

産業財産権の名称 Acoustic signal encoding method, acoustic signal decoding method, program, encoding device, acoustic system, and decoding device	発明者 西口正之、加藤巧大	権利者 同左
産業財産権の種類、番号 特許、17/432098	出願年 2021年	国内・外国の別 外国

産業財産権の名称 音響信号符号化方法、音響信号復号化方法、プログラム、符号化装置、音響システム、及び複合化装置	発明者 西口正之、加藤巧大	権利者 秋田県立大学
産業財産権の種類、番号 特許、PCT/JP2020/006211	出願年 2020年	国内・外国の別 外国

〔取得〕 計1件

産業財産権の名称 音響信号符号化方法、音響信号復号化方法、プログラム、符号化装置、音響システム、及び復号化装置	発明者 西口、加藤	権利者 同左
産業財産権の種類、番号 特許、7232546	取得年 2023年	国内・外国の別 国内

〔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	渡辺 貫治 (Watanabe Kanji) (20452998)	秋田県立大学・システム科学技術学部・准教授  (21401)	

## 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
---------	---------