

様々な音を聞き分けるコンピュータの耳を作るための多チャネル音響信号処理の研究

代表研究者 牧 野 昭 二 早稲田大学 大学院情報生産システム研究科 教授

1 はじめに

都会の街角に立っていると、目を閉じたままでも、車が通り過ぎる音、人々が話す音、通り過ぎる足音、降り続く雨音など、次々と聞こえてくる音を認識することができる。これらすべての音を認識し、知覚された光景を街の音響シーンとして認識することは、人間にとって自然なことである。人間の聴覚は、音源を分別し、関心のある音源に注意を向けることに優れている。この現象は、騒がしい部屋で一つの会話に集中できることになぞらえて、カクテルパーティー効果と呼ばれている。

「音」は豊富な情報を伝達する重要なメディアである。人間は日常的に様々な音を聞くことにより周囲で何が起きているかを知覚し、行動に役立てている。人間の周りに存在する音声に限らない様々な音を認識できるコンピュータの耳をロボットに付ければ、例えばペットの鳴き声を認識してリアリティのある会話の実現でき、車に付けば例えばクラクションの音を認識して自動運転の安全性が増し、家に付けば例えば窓が割れる音を認識してホームセキュリティが向上する。このような様々な音を聞き分ける人間の聞き取り能力をコンピュータに付与することにより、人間のために役に立つ機械（ロボット、車、家、等）を作る意義は大きい。

音響信号処理の分野では、長い間音声が最も重要な対象であったが、近年の音響信号処理の発展に伴い、その対象は音声から音楽、そしてあらゆる音へと拡大してきた。あらゆる音を対象とし、音から環境を解析・理解するための基盤技術は、スマートセキュリティ、高齢者見守り、ロボット聴覚、などの様々な新しい音情報サービスを可能にする技術と期待されている。また、近年では競争型ワークショップ[Detection and Classification of Acoustic Scenes and Events 2025]や国際会議でのスペシャルセッション[Acoustic Event Detection and Classification, ICASSP2025] 等が毎年開催されるなど、音響信号処理分野で、音環境の解析と理解は大変な注目を集めている研究トピックである。

「音」を認識する研究は、深層学習を始めとする機械学習技術の急速な発展を背景に、近年、ますます活発化している。音から環境を解析・理解するための基盤技術に関して、2016 年より Detection and Classification of Acoustic Scene and Event (DCASE) という国際コンテスト（チャレンジ）、ならびに国際会議（ワークショップ）が毎年開催され参加者が急増している。この研究分野の新しい潮流は、IEEE にも大きな影響を与え、IEEE Signal Processing Society (SPS) は新しい EDICS カテゴリーを採用した。2024 年のフラグシップ国際会議 ICASSP では、このカテゴリーに 100 件以上の投稿があった。音響イベント検出は、企業活動においても重要な位置を占めている。故障診断や事故防止のための音によるモニタリングや監視のアプリケーションなど、音響イベント検出のみに特化した企業も数多く存在する。産業上の重要性は、例えば DCASE のワークショップへの企業の幅広い参加によって示されており、ワークショップ参加者の約半数は企業である。また、当該分野のトップジャーナルである IEEE/ACM Trans. Audio Speech and Language Processing 誌でも 2017 年に Sound Scene and Event Analysis という特集号が企画され、40 編を超える投稿を集めたことなどからも、その盛り上がりがわかる。こうした「音」を認識する研究は、深層学習を始めとする機械学習技術の急速な発展を背景に、近年、ますます活発化している。そして、「音」の認識においては、認識方法としての機械学習とともに、どのように音情報を取得するか、すなわち音のセンシング方法もまた、極めて重要である。

実際、音響シーン認識・音響イベント検出では、対象音はマイクロホンから遠く離れていることが多いため、マイクロホンで受信されたときの音圧レベルは、環境中で発生する他の音の音圧レベルよりも低くなる 경우가多く、検出の難易度が増す。また、自然環境には多くの音が存在し、複数の音が同時に存在することがある。雑音、残響、音声の重なりなど、独立に発生する複数の音響イベント同士の重なりをどう扱うかが音響シーン認識・音響イベント検出における重要なポイントとなっている。

2 多チャネル音響信号処理の研究

2-1 音源分離・雑音抑圧・残響除去の統合および移動音源への対応

これまで音源分離・雑音抑圧・残響除去は、異なるアプローチで個別に発展してきており、音源分離の性能は、雑音、残響の影響により劣化し、音源分離における雑音・残響耐性の向上が期待されていた。また、従来妨害音が移動する場合には、妨害音方向へ1つの null（指向性の死角）を形成し、その方向を適応的に変化させることにより対応していた。そのため、適応アルゴリズムの収束速度と妨害音の移動速度の関係に応じて追従性能が決まっていた。そのため、速く移動する妨害音に追従することはできなかった。

本研究では、これまで異なるアプローチで個別に発展してきた音源分離・雑音抑圧・残響除去を同一の誤差関数の最小化問題として統合し、同時最適化する統合手法を開発した。更に、演算量を削減しながら、性能を最適化するための実時間動作実現手法も検討した。

従来、妨害音が移動する場合には、妨害音方向へ1つの null（指向性の死角）を形成し、その方向を適応的に変化させることにより対応していた。そのため、適応アルゴリズムの収束速度と妨害音の移動速度の関係に応じて、追従性能が決まっていた。そのため、速く移動する妨害音に追従することはできなかった。これに対して、本研究では、移動音源を1つの null でカバーできる範囲（角度 15 度程度）のミニブロックごとに捉え、複数の null を同時に形成することにより移動範囲全体をカバーする新方式を提案した。そのために、過去のフィルタ係数を幾何学的制約として与えて過去のフィルタ係数を保持しながら新しいフィルタ係数を求める手法を検討した。

（1）DiffCBF：拡散モデルと畳み込みビームフォーマを用いた音声分離・雑音除去・残響抑圧の統合型音声処理手法の研究

本研究では、雑音除去、残響除去、および音源分離を同時に行う新しいマルチチャネル音声強調手法を提案した。提案手法は、拡散モデルに基づく手法であるマルチストリーム Score-based Generative Model for Speech Enhancement (ms-SGMSE) と、信号処理に基づく手法である畳み込みビームフォーマ (CBF) を組み合わせたものであり、この統合手法を Diffusion model with CBF (DiffCBF) と呼ぶ(図1)。両手法の利点を活用することで、反復的な精緻化により推定精度を向上させた。

さらに、提案手法は音源数が与えられていれば、その数に依らず観測信号を処理可能であり、高い汎用性を有している。実験の結果、提案手法は雑音および残響を含む音声混合からクリーン音声の品質を大幅に向上させ、従来の拡散モデルベースの音源分離手法を大きく上回る性能を示した[1]。

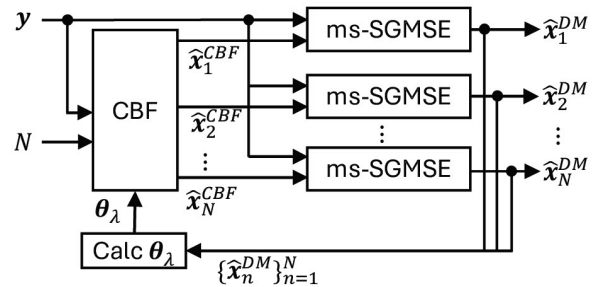


図1 DiffCBF(Diffusion model with CBF)の処理フロー

（2）拡散雑音環境下における高速ブラインド音源抽出および残響除去のための独立低ランク行列分析の研究

本研究では、残響および拡散雑音環境におけるブラインド音源抽出問題を扱った。従来の拡散雑音対応独立低ランク行列分析 (NoisyILRMA) は、拡散雑音環境において効率的な抽出を実現していたが、残響環境では性能が低下していた。

この課題に対処するために、本研究では従来手法に重み付き予測誤差法に基づく残響除去モジュールを統合した NoisyILRMA-conv を提案した(図2)。提案手法は、統一的な最適化枠組みの下で音源抽出と残響除去を同時に実現しつつ、従来手法の計算効率を維持することを可能にした[2]。

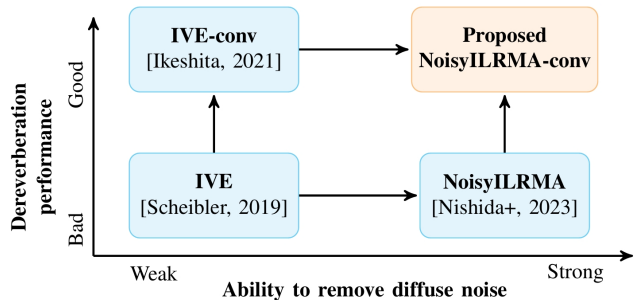


図2 提案手法と従来手法の関係

(3) 一次反射モデルを用いたニューラルネットワーク支援型音源到来方向推定およびビームフォーミングの統合手法の研究

音源信号抽出および到来方向推定は、マイクロホンアレイ信号処理における中核的な課題である。しかし、残響環境下ではこれらの性能が大きく低下していた。すなわち、到来方向推定は強い反射の中から直接音を識別することが困難であり、音源抽出は初期反射に含まれる空間情報を十分に活用できないという問題があった。

本研究では、音源の伝搬特性を、直接音および周囲の壁からの一次反射に対応する自由音場伝搬特性の線形結合として表現する新しい定式化を提案した。さらに、到来方向およびこれらの伝搬特性成分の結合重みを同時に推定するニューラルネットワークを構築した。得られた合成伝搬特性を最小分散無歪応答ビームフォーマに適用することで、音源抽出性能を向上させた(図3)。

加えて、直接音の方向を自動的に識別することを可能にする位置ベースの学習戦略を導入し、エンドツーエンド学習を実現した。シミュレーション実験の結果、提案手法は残響環境において音声品質および到来方向推定精度の両方を大幅に改善することを示した[3]。

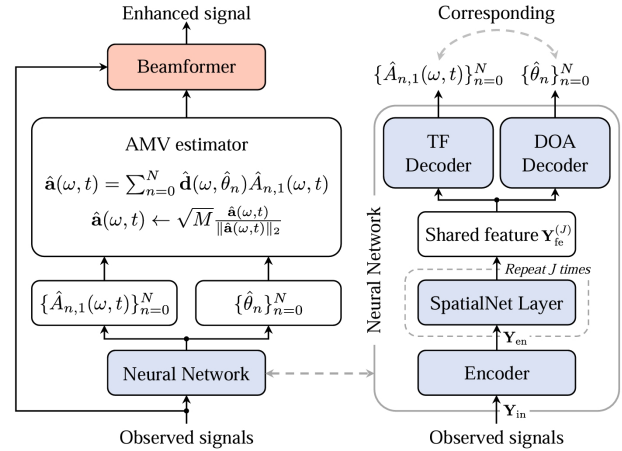


図3 提案するニューラル推定器の構成図

(4) 拡反復的音源ステアリングを用いた高速畳込み伝達関数に基づく多チャンネルNMFの研究

数多くのブラインド音源分離手法の中でも、畳込み伝達関数に基づくマルチチャンネル非負値行列因子分解は、遅延した音源信号の複数フレーム間の相関をモデル化することで、強い残響環境において優れた性能を示していた。しかしながら、各音源ごとに逆行列演算を必要とする反復射影更新則に伴う高い計算コストが、実用展開の妨げとなっていた。

この問題に対処するために、本研究では分離フィルタの更新において逆行列演算を必要としない更新則である反復音源ステアリングを統合した効率的な手法を提案した。実験の結果、提案手法は従来の畳込み伝達関数に基づくマルチチャンネル非負値行列因子分解と同等またはそれ以上の分離性能を達成しつつ、計算量を大幅に削減できることを示した(図4)[4]。

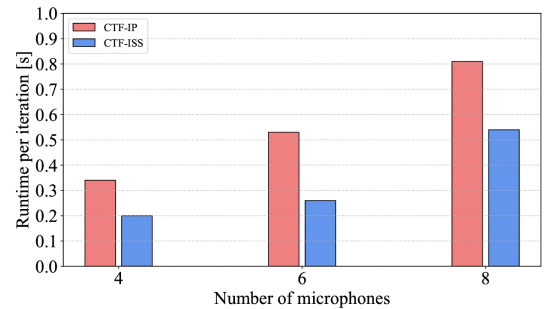


図4 異なるマイクロホン構成における平均処理時間

(5) DySiME: 時間変動する音源到来方向情報を用いた動的単一音源多チャンネル音声強調の研究

単一音源マルチチャンネル音声強調は、雑音および残響環境下において所望の目的話者の音声抽出しつつ、マルチチャンネル出力における空間的整合性を維持することを目的としていた。この問題に対する近年の手法であるマルチチャンネル音源抽出は、方向および時間情報を活用し、固定音源を仮定した残響環境においてエンドツーエンドの複素スペクトログラムマッピングを実現していた。しかし、この手法は事前知識や手動設計による循環的位置埋め込みに強く依存しており、実環境での適用性に課題があった。

これらの制約を克服するために、本研究では目的話者の移動に対応したエンドツーエンドフレームワークであるDySiME (dynamic single-source multichannel enhancement) を提案した(図5)。提案手法は、広帯域および狭帯域の融合に基づく音源定位による到来方向推定モジュールを統合し、目的音源の方向を継続的に追跡した。さらに、学習可能な位置情報アダプタにより、到来方向推定器の中間特徴を強調モデルのバックボーンに組み込むことで、時間的に変動する空間情報を活用した高精度な音声強調を可能にした。

この設計により、推論時における事前の到来方向情報への依存を低減し、移動音源に対しても頑健かつリアルタイムな音声強調を実現した。4 チャンネル円形マイクロホンアレイを用いたシミュレーション評価の結果、提案手法は音声品質および空間精度の両面においてベースラインを一貫して上回る性能を示した[5]。

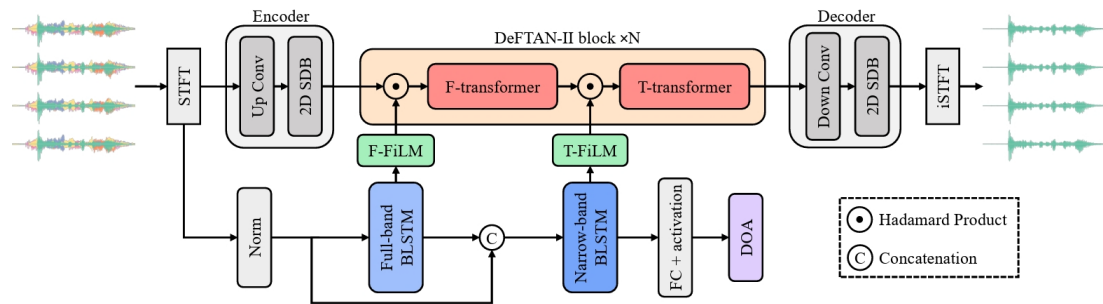


図5 提案する DySiME (dynamic single-source multichannel enhancement) の構成

(6) 再学習不要の拡散モデルにおける U-Net 構成要素の戦略的再重み付けによる音声強調性能の向上の研究

U-Net アーキテクチャを活用した拡散モデルベースの音声強調は、その優れた性能により大きな注目を集めている。本研究では、音声強調タスクにおける拡散モデル内の U-Net の本質的な潜在能力を探索し、拡散過程における音声品質を大幅に向上できることを示した。

音声強調の雑音除去過程における U-Net 各構成要素の役割を分析した結果、主幹ネットワーク（バックボーン）は主に雑音除去に寄与し、スキップ接続は主として高周波成分をデコーダに導入する役割を担うことを明らかにした。しかし、このスキップ接続の効果により、バックボーンが保持する重要な意味情報が相対的に軽視される可能性があった。

この知見に基づき、本研究では再学習やファインチューニングを必要としない手法を提案した。具体的には、U-Net のスキップ接続とバックボーンの特徴マップの寄与を戦略的に再重み付けすることを既存の拡散モデルに容易に組み込み(図6)、音声強調性能の向上を実現した[6]。

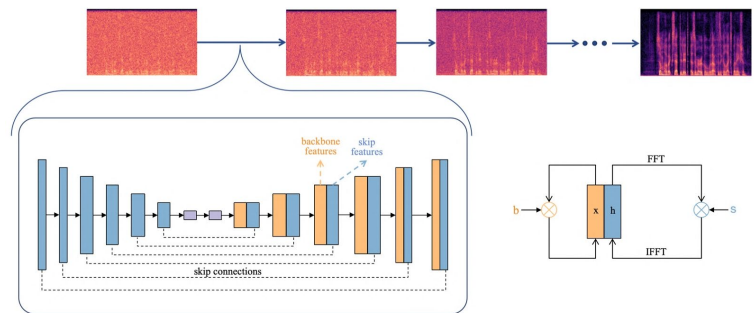


図6 提案手法のフレームワーク

2-2 時間周波数スイッチングフィルタの複素フィルタとしての新展開

線形ビームフォーマに基づく音声強調は、強力だが劣決定条件下（マイク数 $M < \text{音源数 } N$ ）において性能が劣化する。近年普及しているスマートフォンなど、少ないマイクを持つ小型機器でも効果的な音声強調を達成する手法の開発が求められている。

2マイクのビームフォーマでは1個の妨害音を抑圧できる。そこで、我々はこれまでに、 $(N-1)$ 個の妨害音をそれぞれ抑圧する $(N-1)$ 個のビームフォーマを構成し、それらの係数を最小値選択により時間周波数点ごとにスイッチングすることにより、音声強調を行う時間周波数スイッチングビームフォーマの基本コンセプトを提案した。一方、音源分離・雑音抑圧・残響除去は優決定条件下（マイク数 $M > \text{音源数 } N$ ）においてしか動作せず、劣決定条件下において動作する手法が切望されていた。

時間周波数スイッチングビームフォーマは、複素フィルタを用いてマイク間の位相差を制御して空間指向性を形成していた。この概念を音源分離・雑音抑圧・残響除去に展開した時間周波数スイッチングフィルタにより、音源数に依存することなく高品質な出力を得るための統一的なアレイ信号処理を検討した。

時間周波数スイッチングフィルタは、我々らのオリジナル技術であり、音源分離・雑音抑圧・残響除去の統合手法と融合させ、優決定問題（音源数 < マイク数）と劣決定問題（マイク数 < 音源数）の両方に対応できる手法を構築した。

（７）幾何的制約付き移動音源抽出のためのスイッチング時不変分離ベクトルの研究

目的音源が移動する場合、ブラインド音源抽出は特に困難であった。時不変分離ベクトルに基づく独立ベクトル抽出は、ブロックごとに変化する混合モデルを用いながら時不変の分離ベクトルを推定することで移動音源の抽出を行っていたが、マイクロホン数が限られる場合には性能が低下していた。

この制約を克服するために、時不変分離ベクトルに基づく独立ベクトル抽出では時間周波数ポイントごとに異なる分離ベクトルを選択するスイッチング機構が導入されていた。しかしながら、スイッチング時不変分離ベクトル独立ベクトル抽出は状態間のパーミュテーションに起因する不安定性の問題を抱えていた。

本研究では、幾何制約を導入したスイッチング時不変分離ベクトルに基づく独立ベクトル抽出手法を提案した（図 7）。提案手法では、目的音源の到来方向情報を活用することで、スイッチング分離ベクトルの最適化を誘導し、ロバスト性を向上させた。さらに、最適化過程における数値的安定性を高めるために、状態正則化手法を導入した。

実験の結果、提案手法はさまざまなマイクロホン構成において、抽出性能および安定性の両方を大幅に改善することが示された[7]。

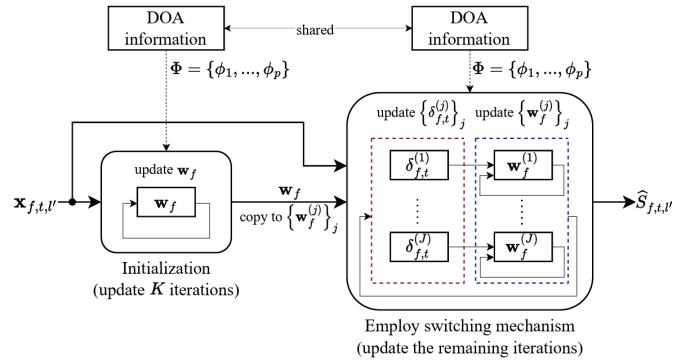


図 7 提案手法の最適化手順のフローチャート

（８）切替型残響抑圧と空間正則化独立ベクトル分析の統合手法の研究

スイッチング独立ベクトル分析は、複数の分離行列に対するスイッチング機構を用いることで分離精度の向上を図るブラインド音源分離アルゴリズムである。しかし、音源に関する事前情報が得られない場合には状態間の置換問題が生じるという課題があった。さらに、残響環境下では性能が低下する傾向があった。

本研究では、スイッチング独立ベクトル分析を空間正則化された分離手法を備えた畳込みビームフォーミングアルゴリズムへと拡張した（図 8）。実験により、提案手法は残響環境下において優れた分離性能を達成するとともに、状態間の置換問題に対しても高いロバスト性を維持することが確認された[8]。

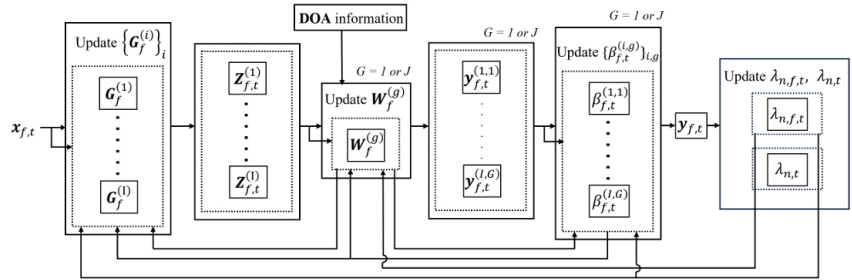


図 8 提案手法の最適化手順のフローチャート

（９）反復的音源指向制御を用いた空間正則化スイッチング独立ベクトル分析の研究

本研究では、反復音源ステアリング更新を組み込んだ空間正則化スイッチング独立ベクトル分析に基づく、計算効率の高い音声分離手法を提案した。提案フレームワークでは、スイッチング機構を導入することで、複数の分離行列を用いてマルチチャネル混合の時間変動特性をモデル化した。また、到来方向情報に基づく空間正則化により、状態間の置換問題を解決した。

さらに、効率向上のために、従来の逆行列演算に基づく更新則を反復音源ステアリング更新に基づくランク 1 更新に置き換えることで、分離性能を維持しつつ計算コストを大幅に削減した（図 9）。

雑音および残響環境における実験の結果、提案手法は従来手法と比較して分離品質の向上および収束の高速化を達成し、限られたマイクロホンアレイにおける実用的な応用に適していることを示した[9]。

2-3 音環境の理解

近年、音響イベント検出における学習データのラベル付けのコストが問題となっている。これを削減するため、我々はBLSTM-CTC (Bi-directional Long Short-Term Memory - Connectionist Temporal Classification) を用いた弱ラベル(音響イベントの時間区間情報を含まない)による学習を開発した。また、複数のビームフォーマとアテンション機構を導入することにより、識別制度を大きく改善できることを確認している。

一方、敵対的学習やマルチタスク学習の枠組みを取り入れることにより、音環境の理解における弱ラベルおよびラベル無し学習を可能にし、ビッグデータのラベル付けコストを大幅に削減する手法が望まれていた。

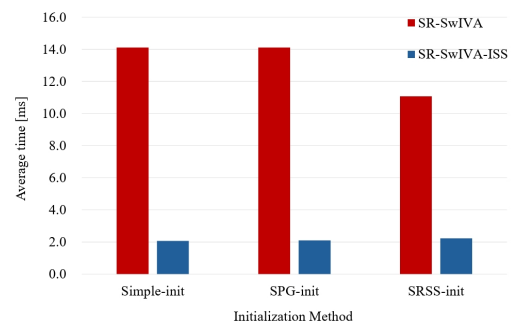


図9 異なる初期化設定における平均処理時間

(10) Mixup と統合したエラスティック加法角度マージン損失による異常音検出の研究

教師なし異常音検知は、正常音のみをモデル化することで異常音を検出することを目的としている。従来手法では、分類ベースの自己教師ありニューラルネットワークモデルを用いて正常音と異常音を識別していた。この考えに基づき、正常クラスのクラス内コンパクト性とクラス間分離性を高めるために加法的角度マージン損失が利用されていた。さらに、従来手法ではミックスアップを導入することでクラス内コンパクト性を確保し、正常データと異常データの分布の重なりを緩和していた。

しかしながら、加法的角度マージンは単一のターゲットロジットにのみ適用され、ミックスアップによって生成された補間ラベルおよびサンプルの学習には適用されないという問題があった。また、固定されたマージンはクラス内およびクラス間のばらつきが不均一なデータには適していなかった。

これらの課題を解決するために、本研究では弾性角度マージン損失をミックスアップデータ拡張に適用した学習手法を提案した(図10)。提案手法では、補間係数に応じてマージンペナルティに重み付けを行うことで、加法的角度マージン損失が補間サンプルに適切に適用できるようにした。さらに、異なる重み付けを持つマージン間で比例的なランダム化を導入することで、クラス内およびクラス間のばらつきの不均一性への対応を強化した。

DCASE 2020 Challenge Task 2 データセットにおける実験の結果、提案手法は AUC 95.03%、pAUC 89.95%を達成し、加法的角度マージン損失とミックスアップを組み合わせた従来手法を上回る性能を示した[10]。

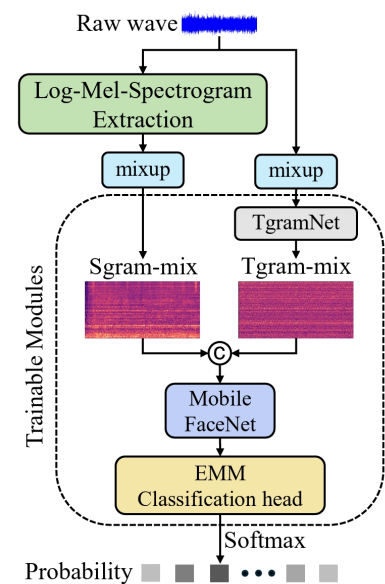


図10 提案する Elastic-Margin-Mixup の学習アーキテクチャ

(11) 音楽分類のためのアテンションベース適応型構造的パッチアウトスペクトログラムトランスフォーマーの研究

音楽分類は音響分類の一分野であり、さまざまな種類の音楽を理解し識別することを目的としている。従来手法は、音響分類のための効率的なトランスフォーマーモデルであり、その中核はトークン削減機構であるパッチアウトであった。パッチアウトは入力系列を短縮することで計算コストを削減し、モデルの汎化性能を向上させていた。しかしながら、そのランダムなパッチ選択は重要な情報の損失を引き起こす可能性が

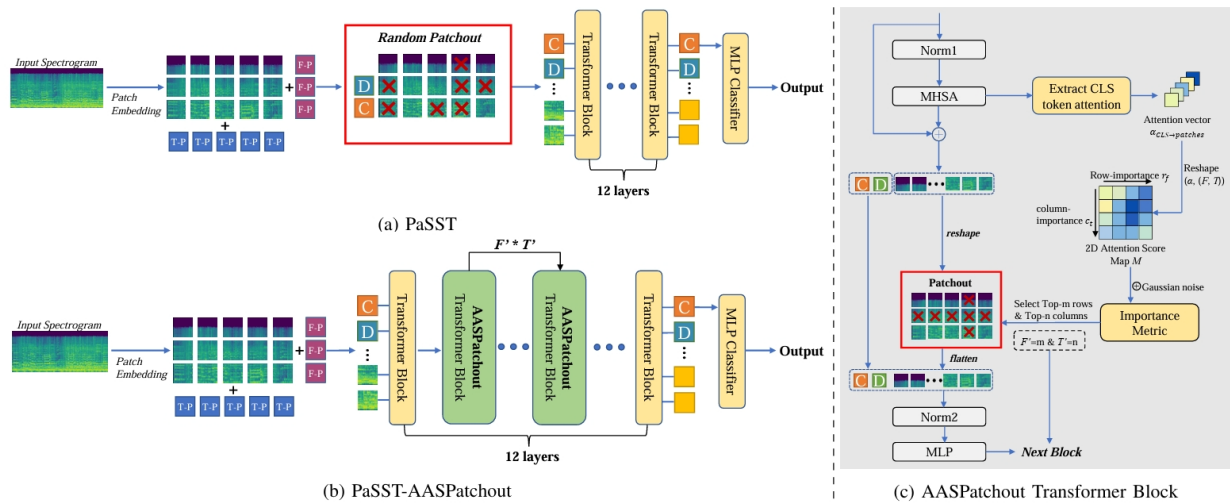


図 11 全体アーキテクチャの比較。(a) PaSST のアーキテクチャ、(b) 提案する PaSST-AASPatchout のアーキテクチャ、(c) AASPatchout Transformer Block の詳細構造

あり、特に構造的に豊かな音楽において特徴学習を十分に行えないという問題があった。

この問題に対して、本研究では新たにアテンションに基づく適応的構造化パッチアウトを提案した(図 11)。提案手法は、トランスフォーマ内部のアテンションスコアを用いることで、最も情報量の多いパッチを保持する。提案手法は追加のモジュールを必要とせず、既存のアテンション機構を直接活用する点に特徴がある。

音楽分類における 2 つのタスク、すなわち多声音楽器認識および音楽ジャンル分類において評価を行った結果、提案手法は従来のランダムパッチアウトベースラインを上回る性能を示した[11]。

3 おわりに

本研究調査では、様々な音を聞き分けるコンピュータの耳を作るための多チャンネル音響信号処理の研究調査として、1) 音源分離・雑音抑圧・残響除去の統合および移動音源への対応、2) 時間周波数スイッチングフィルタの複素フィルタとしての新展開、3) 音環境の理解 に関する検討を行い、国際会議発表 11 件の研究成果を得た。

【参考文献】

- [1] R. Kimura, T. Ueda, T. Nakatani, N. Kamo, M. Delcroix, S. Araki, and S. Makino, "DiffCBF: A diffusion model with convolutional beamformer for joint speech separation, denoising, and dereverberation, " Invited in Proc. EUSIPCO, pp. 945-949, Sept. 2025.
- [2] S. Feng, L. Zhang, and S. Makino, "Independent low-rank matrix analysis for fast joint blind source extraction and dereverberation in diffuse noise environments, " in Proc. IEEE Global Conference on Consumer Electronics (GCCE), pp. 1-2, Sept. 2025.
- [3] Y. Yang, C. Pan, Q. Gao, J. Benesty, S. Makino, and J. Chen, "Neural network-assisted joint DOA estimation and beamforming with first-order reflection modeling, " in Proc. APSIPA, pp. 19-23, Oct. 2025.
- [4] X. Xie, X. Wang, L. Zhang, Y. Yang and S. Makino, "Accelerated convolutive transfer function-based multichannel NMF using iterative source steering, " in Proc. APSIPA, pp. 149-153, Oct. 2025.

- [5] H. Liang, Y. Yang, X. Zhang, S. Makino, and J. Chen, "DySiME: Dynamic single-source multichannel enhancement using time-varying directional cues, " in Proc. APSIPA, pp. 154-159, Oct. 2025.
- [6] Y. Zhang, Y. Li, Y. Zhao, and S. Makino, "Strategic re-weighting of U-Net components in diffusion models for enhanced speech enhancement without retraining, " in Proc. APSIPA, pp. 819-824, Oct. 2025.
- [7] C. Chen, Y. Yang, Y. Zhao, S. Makino, and J. Chen, "Switching constant separating vector for moving source extraction with geometric constraints, " Invited in Proc. APSIPA, pp. 13-18, Oct. 2025.
- [8] G. Zheng, W. Liu, C. Chen, and S. Makino, "On joint switching dereverberation and spatially regularized independent vector analysis, " in Proc. RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP), pp. 49-52, Feb. 2026. **[Student Paper Award]**
- [9] H. Dong, W. Liu, X. Xie, and S. Makino, "Spatially-regularized switching independent vector analysis with iterative source steering, " in Proc. RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP), pp. 45-48, Feb. 2026. **[Student Paper Award]**
- [10] Y. Zhao, Y. Yang, X. Zhang, and S. Makino, "Elastic additive angular margin loss integrated with mixup for anomalous sound detection, " in Proc. APSIPA, pp. 477-482, Oct. 2025.
- [11] Y. Liu, L. Liu, Y. Yang, and S. Makino, "Attention-based adaptive structured patchout spectrogram transformer for music classification, " in Proc. APSIPA, pp. 282-287, Oct. 2025.

〈 発 表 資 料 〉

題 名	掲載誌・学会名等	発表年月
DiffCBF: A diffusion model with convolutional beamformer for joint speech separation, denoising, and dereverberation	European Signal Processing Conference (EUSIPCO)	September 2025
Independent low-rank matrix analysis for fast joint blind source extraction and dereverberation in diffuse noise environments	IEEE Global Conference on Consumer Electronics (GCCE)	September 2025
Neural network-assisted joint DOA estimation and beamforming with first-order reflection modeling	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025
Accelerated convolutive transfer function-based multichannel NMF using iterative source steering	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025
DySiME: Dynamic single-source multichannel enhancement using time-varying directional cues	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025

Strategic re-weighting of U-Net components in diffusion models for enhanced speech enhancement without retraining	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025
Switching constant separating vector for moving source extraction with geometric constraints	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025
On joint switching dereverberation and spatially regularized independent vector analysis [Student Paper Award]	RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP)	February 2026
Spatially-regularized switching independent vector analysis with iterative source steering [Student Paper Award]	RISP International Workshop on Nonlinear Circuits, Communications and Signal Processing (NCSP)	February 2026
Elastic additive angular margin loss integrated with mixup for anomalous sound detection	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025
Attention-based adaptive structured patchout spectrogram transformer for music classification	Asia Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA ASC)	October 2025