

非同期分散マイクロフォンアレーによる音源定位・音源分離

Source Localization and Separation with Asynchronous and Distributed Microphone Array

小野 順貴

Nobutaka ONO

国立情報学研究所 / 総合研究大学院大学

National Institute of Informatics / The Graduate University for Advanced Studies (SOKENDAI)

onono@nii.ac.jp

Abstract

本稿では、複数の録音機器を分散配置させ、これらをマイクロフォンアレーの素子として用いる、非同期分散マイクロフォンアレーという枠組みについて紹介する。従来のマイクロフォンアレー信号処理においては、チャンネル間の微小な時間差が空間情報の大きな手がかりであり、各チャンネルを正確に同期させるために、全てのマイクロフォンは多チャンネル A/D 変換器に接続されることが前提であった。これに対し、我々の身の周りには、ラップトップ PC、ボイスレコーダー、スマートフォンなどの録音機器が多数存在する。こうした機器によりマイクロフォンアレー信号処理が可能になれば、その利便性は大きく、適用範囲を格段に広げることが期待できる。本稿では、非同期録音機器を用いたマイクロフォンアレー信号処理の新しい展開について、関連研究を概観しつつ、著者らの取り組みを紹介する。

1 はじめに

マイクロフォンアレーは、複数のマイクロフォンにより音場の空間的な情報を取得し、単一マイクロフォンでは困難な、音源定位、音源強調、音源分離などを行う枠組みである。一般には、用いられるマイクロフォンの数が多いほど得られる空間情報が多くなるため、制御できる指向性の自由度が増加し、また、マイクロフォンを広範囲に配置することができるほどカバーできる範囲が広がり、定位や分離の性能向上が期待できる。

しかしながら、マイクロフォンアレー信号処理においては、厳密な同期録音が必要不可欠であることが大きな制約条件となっている。これは、マイクロフォンアレー信号処理では、各マイクロフォンで録音される信号間の微小

な時間差（例えば、経路長 3.4 cm に対して $100 \mu\text{s}$ ）が音源の空間情報の主要な手がかりとなっているためである。よって従来は、各チャンネルを正確に同期させるために、全てのチャンネルは多チャンネル A/D 変換器に接続され、同一クロックによりサンプリングされる必要があり、これがマイクロフォンアレーの多素子化や分散配置などに対して大きなコストを生じる主要因の一つとなっていた。

一方、我々の身の周りには、録音機能を持つ機器が多数存在している。音を録音することが目的であるボイスレコーダーや、通話を目的としたスマートフォンにとどまらず、ラップトップ PC やタブレット端末の多くも録音機能を有しているし、動画撮影機能を持つデジタルカメラやビデオカメラも録音機器として用いることができる。こうした別々の録音機器による多チャンネル録音に基づくアレー信号処理の枠組みは、近年、非同期分散マイクロフォンアレー、アドホックマイクロフォンアレーなどと呼ばれ、国内外で関心が高まりつつある [1]。

非同期分散録音機器でアレー信号処理が可能になれば、以下のような利点が期待できる。

1. 従来のように、同時サンプリング可能な多チャンネル A/D 変換器が不要なため、録音機器を増やすだけで、マイクロフォンアレーの素子数を容易に増やすことができる。
2. 広範囲に素子を分散配置することができる。
3. 録音機器間の有線接続の必要がなく、ワイヤレスのシステムを容易に構築することができる。

本稿では、非同期分散マイクロフォンアレーの新しい展開について、関連研究と共に著者らの取り組みを紹介する。なお本稿では特に、複数の機器を用いて録音した多チャンネル信号に対する信号処理を想定し、全チャンネルの信号が利用可能であることを仮定した音源定位と音源分離の問題を中心に紹介する。著者らの別の解説論文 [2, 3, 4] も併せて参照いただきたい。

また、本稿では扱わないが、マイクロフォン間の通信を仮定し、音のセンサネットワークを構築することを目指すワイヤレスアコースティックセンサネットワークという枠組みでは、各マイクロフォンは近傍のマイクロフォンとした通信できない、つまり全チャンネルの情報が利用できるわけではないという制約の元で、どのように信号処理を行うか、といった問題も盛んに議論されている。興味のある読者は [1] や国際会議 EUSIP2013 のチュートリアル資料 ([5] から参照可) などを参照していただきたい。

2 音源定位

2.1 マイクロフォン位置の自己校正の必要性

従来のアレー信号処理に用いられるマイクロフォンは、直線、円周のような規則的な配置に並べられ、各マイクロフォンの位置は既知であることが前提である。しかしながら非同期分散マイクロフォンアレーでは、個々の録音機器は有線接続すらされておらず、通常、その位置座標は未知であることが多い。そのため、音源位置を推定するためには、まずマイクロフォン位置を推定することが必要となる。スマートフォンなどを用いる場合には GPS (Global Positioning System) 情報が利用できる場合もあるが、室内環境での測位精度は一般に十分でなく、マイクロフォンが観測された音自体を用いて推定する自己校正 (self-calibration) が重要となる。

なお、位置推定は一般にアレー信号処理への応用に限らず重要であり、Indoor Positioning and Indoor Navigation (IPIN)[6] という室内での機器位置推定に関する国際会議が 2010 年以降毎年開催され、音に限らず、電波や光を含めた定位手法が議論されている。

2.2 ブラインドアラインメント

マイクロフォン位置や音源位置を推定するための観測量としては、到来時間 (Time of Arrival; TOA) や到来時間差 (Time Difference of Arrival; TDOA) がよく用いられる。ここでは録音機器が同期しておらず、音源も未知な場合、各音源に対して観測された TDOA のみから音源位置、マイクロフォン位置、各マイクロフォンの時間原点を推定するという問題を考えてみよう。我々はこれを、観測信号のみから (ブラインド)、音源、マイクロフォン、チャンネルを共通の時空間座標系に揃える (アラインメント) という意味で、ブラインドアラインメントと呼んでいる [7, 9]。

M 個のマイクロフォン、 N 個の音源、の位置座標をそれぞれ $r_1, r_2, \dots, r_M, s_1, s_2, \dots, s_N$ とする。ここでは r_m, s_n は p 次元のベクトルとし、 $p = 2$ または $p = 3$ とする。また、各マイクロフォンはそれぞれ別の時間原点をもっており、それらを t_m で表わす。簡単のため、録音機器間のサンプリング周波数 mismatches は、ここでは無視できると仮定する。

まず、音源 s_n に対する TDOA を求めることを考える。チャンネルが非同期である場合、正しい TDOA は直接には求まらず、マイクロフォン i, j の相互相関ピークから得られるチャンネル間の時間差 (みかけの TDOA) は、

$$\hat{\tau}_{nij} = \left(\frac{\|s_n - r_i\|_2}{c} - t_i \right) - \left(\frac{\|s_n - r_j\|_2}{c} - t_j \right) \quad (1)$$

のように、未知の録音開始時刻 t_i, t_j を含むことになる。よって、みかけの TDOA が得られたとしても、これが到来時間差によるものなのか、マイクロフォンの時間原点差によるものなのかはわからない。しかし、

$$(K - p - 1)(M - p - 1) \geq \frac{p(p+3)}{2} \quad (2)$$

の場合には、未知数の数より多くの観測量が得られる [7]。よって、TDOA に基づくチャンネル同期とマイクロフォン位置定位の問題は、式 (1) を観測方程式とし、多数得られた $\hat{\tau}_{nij}$ から s_n, r_m, t_m を推定する問題と考えることができる。我々は、補助関数法 [8] という最適化のアプローチを用いて式 (1) の最小二乗解を求める解法 [7, 9] を導出している。また、近年ロボットへの応用を想定したオンラインでのキャリブレーション [10] も検討されている。

2.3 距離行列のもつランク制約

チャンネルが同期していたとしても、マイクロフォンと音源の同時位置推定は難しい問題である。音速を既知とすれば TOA や TDOA は距離情報を与える。一般に距離情報から座標を推定する問題は、多次元尺度法 (Multi Dimensional Scaling; MDS) という手法によって解くことができるが、マイクロフォンと音源の位置推定問題の場合には、マイクロフォン間、音源間の距離情報は得ることができず、マイクロフォンと音源の間の距離情報しか得られないことが難しさの一因となっている。ここでは、こうした問題を解く上で大変有用な、距離行列の性質について触れておきたい。

いま、距離行列 $D = (D_{mn})_{M \times N}$ 、距離差行列 $\tilde{D} = (\tilde{D}_{mn})_{(M-1) \times (N-1)}$ を、マイクロフォンと音源の距離の 2 乗を用いて以下のように定義する。

$$\begin{aligned} D_{mn} &= \|r_m - s_n\|_2^2 = (r_m - s_n)^T (r_m - s_n) \\ &= \|r_m\|_2^2 + \|s_n\|_2^2 - 2r_m \cdot s_n \end{aligned} \quad (3)$$

$$\tilde{D}_{mn} = D_{m+1, n+1} - D_{m+1, 1} - D_{1, n+1} + D_{1, 1} \quad (4)$$

このとき以下が成り立つ。

補題 1 距離行列 D のランクは高々 $(p+2)$ である。

証明: 式 (3) より、

$$D = T^T I_N + I_M^T U - 2R^T S \quad (5)$$

とあらわせる。ただし、 T は m 番目の要素が $\|r_m\|_2^2$ である $1 \times M$ のベクトル、 U は n 番目の要素が $\|s_n\|_2^2$ である $1 \times N$

のベクトル, I_M, I_N は要素が全て1の $1 \times M, 1 \times N$ のベクトル, $R = (r_1 \cdots r_M)$ は $p \times M$ 行列, $S = (s_1 \cdots s_N)$ は $p \times N$ 行列であり, また, T は行列の転置を表す。各項のランクは高々, $1, 1, p$ であるから, D のランクは高々 $(p+2)$ である。■

補題 2 距離差行列 \tilde{D} のランクは高々 p である。

証明: $\tilde{R} = (r_2 - r_1 \cdots r_M - r_1)$, $\tilde{S} = (s_2 - s_1 \cdots s_N - s_1)$ とおくと, $\tilde{D} = -2\tilde{R}\tilde{S}^T$ とあらわされる。 \tilde{R} は $p \times (M-1)$ 行列, \tilde{S} は $p \times (N-1)$ 行列であるので, \tilde{D} のランクは高々 p である。■

任意の配置に対して, 距離行列がこのようなランク制約をもつことは興味深い。近年, このランク制約を用いた TOA, TDOA ベースの位置推定法が提案されている [11, 12, 13]。我々も現在, いくつかのアルゴリズムを研究中 [14, 15] であり, 今後はブラインドアライメントへ応用していきたいと考えている。

2.4 音の発信を利用した機器位置推定

スマートフォンなど, 音を発することができる録音機器を利用できる場合には, 音の発信を積極的に利用するアプローチが考えられる。例えば, 1 台のスマートフォンに装備されているスピーカーとマイクの位置は厳密には異なっているが, これを近似的に等しい ($r_m \simeq s_m$) とし, 2 台のスマートフォンから互いに音を発信して TDOA を求めたとすると,

$$\hat{\tau}_{mn} = -\frac{\|r_m - r_n\|_2}{c} - t_m + t_n \quad (6)$$

$$\hat{\tau}_{nm} = \frac{\|r_n - r_m\|_2}{c} - t_m + t_n \quad (7)$$

と表せるので,

$$\|r_m - r_n\|_2 = \frac{c}{2}(\hat{\tau}_{nm} - \hat{\tau}_{mn}) \quad (8)$$

$$t_m - t_n = -\frac{1}{2}(\hat{\tau}_{mn} + \hat{\tau}_{nm}) \quad (9)$$

のように, 互いの TOA から距離と時間原点の差が直接的に求まる [16, 17, 18]。各機器間の距離が求まれば, あとは前述の多次元尺度法により相対位置を決めることができる。

我々はさらに, 音の発信を利用し, 位置と時間原点のキャリブレーションだけでなく, サンプル周波数 mismatch もあわせて補償する手法を提案している [19]。図 1 に, 4 台の iPod touch による移動音源定位の実験結果例を示す。それぞれから発信された TSP (Time-Stretched Pulse) 信号 [20] を用いて, 各機器の位置, 録音開始時刻の推定, サンプル周波数 mismatch を推定してキャリブレーションを行った後にスピーカーを音源定位したものであり, 移動音源であるスピーカーの定位がうまく行われていることがわかる。

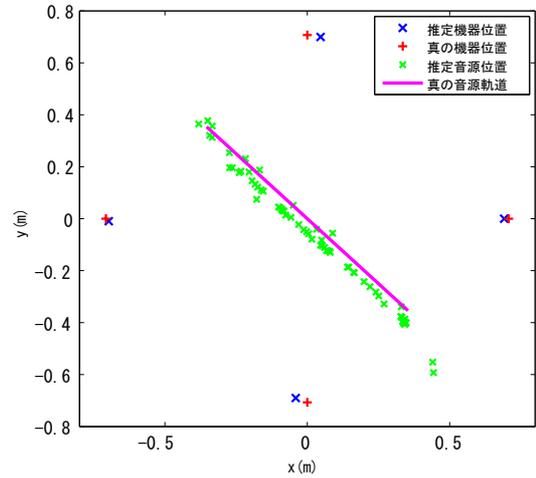
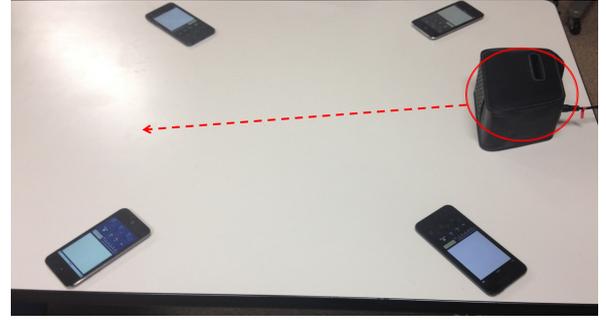


図 1: 非同期分散マイクロフォンアレーによる移動音源定位実験の写真 (上) と実験結果 (下) [19]

3 音源分離

3.1 チャンネル非同期がアレー信号処理に与える影響

一般に非同期分散マイクロフォンアレーにおいては, 通常マイクロフォン位置は未知であり, 各録音チャンネルは同期していない。いわゆるビームフォーミングや音源定位を行うためには, マイクロフォン位置の推定も必要となるが, SN 比最大化ビームフォーマ [21] や独立成分分析 [22] など, 音源強調, 音源分離手法の中には, マイクロフォンの位置情報を必要としない手法もある。そこでここでは主に, チャンネル非同期の影響について考える。

非同期の主な要因には, 1) 録音を開始する時刻が同一でないこと, 2) サンプル周波数が同一でないこと, の 2 つがある。前者は定常的な時間軸シフトを, 後者は時間軸の伸縮をもたらす (図 2 参照)。定常的な時間シフトについては, 信号間の相互相関が最大となるように信号をシフトすることで十分な場合も多い。音源からマイクまでの伝達関数は, SN 比最大化ビームフォーマにおいては学習区間の信号から, 独立成分分析においてはブラインドに推定されるため, たとえ信号間に小さな一定のシフト誤差が残っていても, あまり問題にならないためである。

一方, 録音機器 1, 録音機器 2 のサンプル周波数を f_1, f_2 とすると,

$$\varepsilon = \frac{f_2}{f_1} - 1 \quad (10)$$

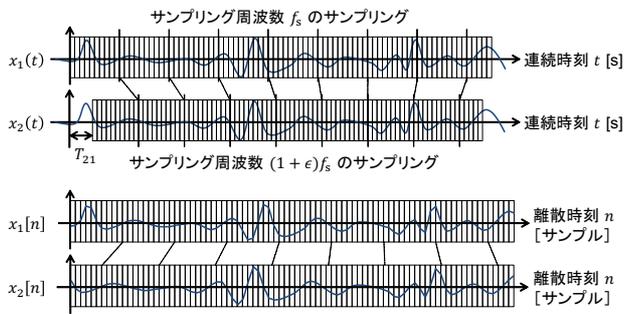


図 2: サンプル周波数ミスマッチの影響：連続波形 (上), 離散波形 (下)

が、録音機器間の相対的なサンプル周波数ミスマッチを表す無次元量となる。 ϵ は、10 ppm (ppm は parts per million で 10^{-6} を表す) の数倍程度に収まることが多い。しかし、こうしたわずかなミスマッチがアレー信号処理に与える影響は大きい。

例えば、録音機器 1, 2 のサンプル周波数を 16000 Hz, 16000.5 Hz とし (サンプル周波数ミスマッチは 31.25 ppm), これらを 0.2 m だけ離れた配置で、正面方向の音源信号を 10 秒間録音を行ったとしよう。簡単のため、録音は同時に開始したと考える。このとき 10 秒間の音響信号は、録音機器 1 では 160000 サンプル、録音機器 2 では 160005 サンプルに相当し、サンプル周波数ミスマッチにより、10 秒間の録音信号の最後では 5 サンプルの時間差が生じる。一方、30 度方向から到来する音波に対する到来時間差は、 $0.2 \sin 30^\circ / 340.0 \times 16000 \approx 4.7$ で、約 4.7 サンプルに相当する。つまり、31.25 ppm という微小なサンプル周波数のずれが引き起こす時間差は、10 秒間に音源が正面から 30 度方向に移動したのと区別がつかないことになる。多くの線形アレー信号処理においては、音源からマイクロフォンまでの伝達関数は線形時不変であることが仮定されているため、こうしたチャンネル間の時間差のドリフトは、音源分離に深刻な破綻を引き起こすことになる [23, 24]。

3.2 サンプル周波数ミスマッチのブラインド補償に基づくブラインド音源分離

サンプル周波数ミスマッチの問題に対しては、ミスマッチを推定し、補償した後に、従来のアレー信号処理を適用する方法が考えられる。前述のように、サンプル周波数ミスマッチは離散波形の伸縮を引き起こすため、同一音源から到来する信号を 2 回観測できれば、それらの間隔の比率からサンプル周波数ミスマッチが推定できる [25] し、観測信号のみからのブラインド推定と補償も検討されつつある [24, 26]。

前述のようにサンプル周波数ミスマッチは、音源が移動しているかのように、時間とともに拡大するチャンネル間時間差を生じる。そこで我々は、音源が動かないこ

とを仮定し、STFT (Short-Time Fourier Transform) 領域で定常な確率モデルに基づく最尤法によりサンプル周波数ミスマッチをブラインドに推定し、補償する手法を提案している [27, 28, 29, 30, 31, 32]。

図 3 は、複数話者の混合音声録音した、複数の長さのデータに対するサンプル周波数の補正精度を平均二乗誤差 (RMSE) で評価した実験結果の例 [28, 29] である。最も短い 3 秒の観測信号でも RMSE は元のサンプル周波数ミスマッチ ϵ の 10 分の 1 以下に収束し、データ長が増えるにつれて急速に小さくなるのがわかる。次に、音源分離への適用例を図 4 に示す。分離フィルタの学習は、補助関数型独立ベクトル分析 [33] により行った。分離性能評価には BSS Toolbox [34] を用い、評価尺度としては SDR (Signal to Distortion Ratio) を用いた。まず未処理の分離性能が非常に低い値を示しており、サンプル周波数ミスマッチの補償をしなければ音源分離ができない厳しい条件であるということが分かる。 ϵ の正しい値を与えた位相補償 (図中位相補償) は、サンプル周波数ミスマッチがない場合よりも SDR が 2 dB 程度低くだけであり、STFT 領域におけるミスマッチ補償法の有効性が確認できる。また、サンプル周波数ミスマッチをブラインドに推定して補償する提案手法は、正しい ϵ を与えた場合とほとんど性能差がなく、提案手法の最尤推定は十分高い精度で ϵ を推定していることを示している。

以上より、非同期録音であっても、音源が動かないと仮定できる場合には、こうした手法によりブラインドにチャンネルを同期化し、ブラインド音源分離が適用できることがわかる。また我々は、符号化録音に対しても、本手法が破綻せず動作することを実験により確認している [35]。

3.3 伝達関数ゲイン基底 NMF を用いた非同期録音に対する音源分離

非同期分散マイクロフォンアレーでは、アレー素子として別々の録音機器を用いるため、関心のある個々の音源の近くにマイクロフォンを配置できる場合も多い。こうした場合にはチャンネル間の時間差だけでなく、チャンネル間の振幅比も音源を区別する重要な手がかりとなる。そこで我々は、サンプル周波数ミスマッチを正確に推定・補償する代わりに、ミスマッチに頑健な振幅スペクトル情報のみを用いた音源分離手法についても検討している。

いま、サンプル周波数ミスマッチの影響により位相情報は信頼できないと考え、通常の線形混合モデルの代わりに、振幅スペクトルに対する近似的な線形混合モデルを仮定する。すなわち、

$$\bar{X}(\tau, \omega) \simeq \bar{A}(\omega) \bar{S}(\tau, \omega) \quad (11)$$

ただし、 $\bar{X}(\tau, \omega)$, $\bar{S}(\tau, \omega)$ は、 $X(\tau, \omega)$, $S(\tau, \omega)$ の要素ごとに絶対値をとったベクトル (振幅スペクトルベクトル)、

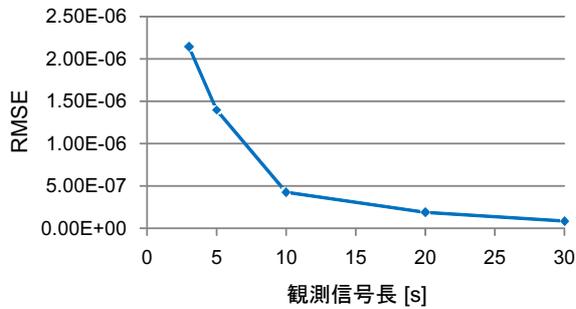


図 3: 推定されたサンプリング周波数ミスマッチ ε の平均二乗誤差 [28]

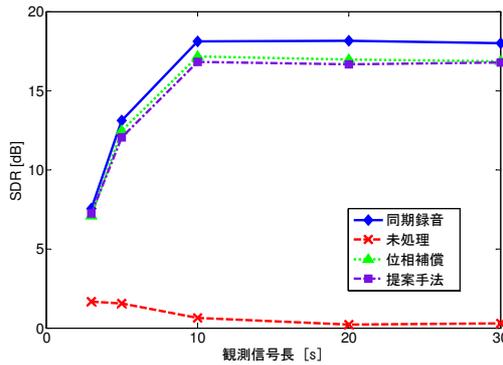


図 4: サンプリング周波数ミスマッチ補償のブラインド音源分離による性能評価 [28, 29]

$\bar{A}(\omega)$ は、各音源から各マイクロフォンまでの伝達関数のゲインを要素とする混合行列である。

このような振幅領域での混合モデルを用いたビームフォーマ [36] も提案されているが、行列が全て非負であることから、非負値行列分解 (Non-negative Matrix Factorization; NMF) の適用も考えられる [37]。ただし、時間周波数領域での NMF と異なり、時間チャンネル領域での NMF の場合には、優決定であったとしても、チャンネル数 (観測行列の行数) と音源数 (基底数) がそれほど大きくは変わらないことが多い。そのような場合には、NMF の低ランク近似としての拘束力が弱く、通常の NMF で観測行列をブラインドに分解することは難しい。

よって我々はまず、各音源に対する単一音源観測が教師信号として得られることを仮定し、あらかじめ $\bar{A}(\omega)$ を学習した後、観測信号 $\bar{X}(\tau, \omega)$ に対して、 $\bar{S}(\tau, \omega)$ のみを更新していく、教師あり NMF の適用を非同期録音に対して検討した。以下では、実際の会議録音を想定した非同期録音データにより、教師あり NMF による音源分離の性能を評価した例 [38] を示す。

分離性能を定量評価するために、まず話者ごとに同期録音を行い、その後、話者ごとの録音を足し合わせて混合した後、マイクごとに、サンプリング周波数が 16000, 16001, 16002, 16003Hz になるようサンプリングを行って、人工的に非同期録音データを生成したものを対象と

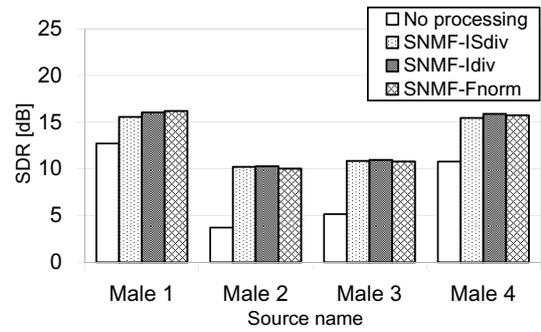


図 5: 非同期録音に対する伝達関数ゲイン基底 NMF の音源分離性能 [38]

した。分離性能評価には前述の BSS Toolbox [34] を用い、評価尺度としては SDR を用いた。また評価は、未処理 (No processing) のほか、教師あり伝達関数ゲイン NMF を板倉斎藤ダイバージェンス規準 (SNMF-ISdiv), I ダイバージェンス規準 (SNMF-Idiv), フロベニウスノルム規準 (SNMF-Fnorm) で適用したものに対して行った。

結果を図 5 に示す。未処理の場合と比較して、教師ありの伝達関数ゲイン基底 NMF では各観測信号において SDR が大きく向上している。また、ここでは結果を省略しているが、同期録音と非同期録音の結果にほとんど差が見られなかった。よってこの手法が非同期録音に頑健であり、非同期分散マイクロフォンアレーにおける目的音強調に適した手法であることが確認できる。我々はその他、マイク数・マイク配置との関係 [39] や、スパース性を導入したブラインド化 [40] についても検討を進めている。

4 おわりに

本稿では、非同期分散マイクロフォンアレーという新しいアレー信号処理の枠組みについて紹介した。十分紹介しきれなかったトピックもあるが、引用した文献をあわせて参照していただければ幸いである。

謝辞

本稿で紹介した研究は、筑波大学牧野研究室、東京大学亀岡研究室との共同研究によるものである。ここに共同研究者各位へ謝意を表す。また、本研究の一部は、科学研究費補助金基盤研究 (B) (課題番号 25280069) の助成を受けて行われたものである。

参考文献

- [1] A. Bertrand, “Applications and trends in wireless acoustic sensor networks: a signal processing perspective,” Proc. Symposium on Communications and Vehicular Technology (SCVT), 2011.
- [2] 小野 順貴, Trung-Kien Le, 宮部 滋樹, 牧野 昭二, “アドホックマイクロフォンアレー – 複数のモバイル録音機器で行う音響信号処理 –,” 電子情報通信学会 Fundamental Review, vol. 7, no. 4, pp. 336–347, 2014.

- [3] https://www.jstage.jst.go.jp/article/essfr/7/4/7_336/_pdf
- [4] 小野 順貴, 宮部 滋樹, 牧野 昭二, “非同期分散マイクロホンアレイに基づく音響信号処理,” 日本音響学会誌, vol. 70, no. 7, pp. 391–396, 2014.
- [5] <http://www.eurasip.org/Proceedings/Eusipco/Eusipco2013/>
- [6] <http://ipin-conference.org/>
- [7] N. Ono, H. Kohno, N. Ito and S. Sagayama, “Blind alignment of asynchronously recorded signals for distributed microphone array,” Proc. WASPAA, pp.161–164, Oct. 2009.
- [8] 小野 順貴, “補助関数法による最適化アルゴリズムとその音響信号処理への応用,” 日本音響学会誌, vol. 68, no. 11, pp. 566–571, 2012.
- [9] K. Hasegawa, N. Ono, S. Miyabe and S. Sagayama, “Blind estimation of locations and time offsets for distributed recording devices,” Proc. LVA/ICA, pp. 57–64, Sept. 2010.
- [10] H. Miura, T. Yoshida, K. Nakamura and K. Nakadai, “SLAM-based online calibration of asynchronous microphone array for robot audition,” Proc. IROS, pp. 524–529, 2011.
- [11] M. Crocco, A. Del Bue and V. Murino, “A bilinear approach to the position self-calibration of multiple sensors,” IEEE Trans. Signal Process., vol. 60, no. 2, pp. 660–673, Feb. 2012.
- [12] F. Jiang, Y. Kuang and Kalle Åström, “Time delay estimation for TDOA self-calibration using truncated nuclear norm regularization,” Proc. ICASSP, pp. 3885–3889, May, 2013.
- [13] Y. Kuang and Kalle Åström, “Stratified sensor network self-calibration from tdoa measurements,” Proc. EUSIPCO, 2013.
- [14] T. -K. Le and N. Ono, “Numerical Formulae for TOA-based Microphone and Source Localization” Proc. IWAENC, pp. 179–183, Sept. 2014.
- [15] T. -K. Le and N. Ono, “Reference-distance Estimation Approach for TDOA-based Source and Sensor Localization,” Proc. ICASSP, 2015. (submitted)
- [16] V. Raykar, I. Kozintsev and R. Lienhart, “Position calibration of microphones and loudspeakers in distributed computing platforms,” IEEE Trans. Speech Audio Process., vol. 13, no. 1, pp. 70–83, 2005.
- [17] M. Hennecke, T. Plötz, G. A. Fink, J. Schmalenböcker and R. Hab-Umbach, “A hierarchical approach to unsupervised shape calibration of microphone array networks,” Proc. SSP, pp. 257–260, 2009.
- [18] 柴田 一暁, 小野 順貴, 亀岡 弘和, “音の発信を利用したスマートフォンアレイの機器位置推定,” 音講論 (秋), pp. 591–592, 9月, 2013年.
- [19] 柴田 一暁, 小野 順貴, 亀岡 弘和, “音の発信を利用したキャリアレーションに基づくアドホックマイクロホンアレイによる音源定位,” 音講論 (春), pp. 707–710, 3月, 2014年.
- [20] Y. Suzuki, F. Asano, H.-Y. Kim, and Toshio Sone, “An optimum computer-generated pulse signal suitable for the measurement of very long impulse responses,” J. Acoust. Soc. Am., vol. 97, no. 2, pp. 1119–1123, 1995.
- [21] H. L. Van Trees, Ed., Optimum Array Processing, John Wiley & Sons, 2002.
- [22] A. Hyvärinen, J. Karhunen and E. Oja, Independent Component Analysis, John Wiley & Sons, 2001.
- [23] E. Robledo-Arnuncio, T. S. Wada and B.-H. Juang, “On dealing with sampling rate mismatches in blind source separation and acoustic echo cancellation,” Proc. WASPAA, pp. 34–37, Oct. 2007.
- [24] Z. Liu, “Sound source separation with distributed microphone arrays in the presence of clock synchronization errors,” Proc. IWAENC, 2008.
- [25] R. Sakanashi, N. Ono, S. Miyabe, T. Yamada and Shoji Makino, “Speech enhancement with ad-hoc microphone array using single source activity,” Proc. APSIPA, Oct. 2013.
- [26] S. Markovich-Golan, S. Gannot and I. Cohen, “Blind sampling rate offset estimation and compensation in wireless acoustic sensor networks with application to beamforming,” Proc. IWAENC, 2012.
- [27] 宮部 滋樹, 小野 順貴, 牧野 昭二, “非同期録音信号の線形位相補償によるブラインド同期と音源分離への応用,” 音講論 (秋), pp. 689–690, 9月, 2012年.
- [28] 宮部 滋樹, 小野 順貴, 牧野 昭二, “非同期録音ブラインド同期のための線形位相補償の効率的な最尤解探索,” 音講論 (春), pp. 733–734, 3月, 2013年.
- [29] S. Miyabe, N. Ono and S. Makino, “Blind compensation of inter-channel sampling frequency mismatch with maximum likelihood estimation in STFT domain,” Proc. ICASSP, pp.674–678, May, 2013.
- [30] 宮部 滋樹, 小野 順貴, 牧野 昭二, “非整数サンプルシフトのフレーム分析を用いた非同期録音の同期化,” 音講論 (秋), pp. 593–596, 9月, 2013年.
- [31] S. Miyabe, N. Ono and S. Makino, “Optimizing frame analysis with non-integer shift for sampling mismatch compensation of long recording,” Proc. WASPAA, Oct. 2013.
- [32] S. Miyabe, N. Ono and S. Makino, “Blind Compensation of Interchannel Sampling Frequency Mismatch for Ad hoc Microphone Array Based on Maximum Likelihood Estimation,” Elsevier Signal Processing (to appear)
- [33] N. Ono, “Stable and fast update rules for independent vector analysis based on auxiliary function technique,” Proc. WASPAA, pp. 189–192, Oct. 2011.
- [34] E. Vincent, C. Fevotte and R. Gribonval, “Performance measurement in blind audio source separation,” IEEE Trans. Acoust. Speech Signal Process., vol. 14, no. 4, pp. 1462–1469, 2006.
- [35] 宮部 滋樹, 小野 順貴, 牧野 昭二, 高橋 祐, “非同期マイクロホンアレイの符号化録音におけるビットレートと同期性能の関係,” 音講論 (春), pp. 725–726, 3月, 2014年.
- [36] 加古 達也, 小林 和則, 大室 伸, “非同期分散マイクロホンアレイのための振幅スペクトルビームフォーマの提案,” 音講論 (春), pp. 829–830, 3月, 2013年.
- [37] 戸上 真人, 川口 洋平, 小窪 浩明, 大淵 康成, “音源のチャンネル間振幅差を基底ベクトルとする音源分離,” 音講論 (春), pp. 803–804, 3月, 2010年.
- [38] 千葉 大将, 小野 順貴, 宮部 滋樹, 山田 武志, 牧野 昭二, 高橋 祐, “伝達関数ゲイン基底 NMF による分散配置非同期録音における目的音強調の検討,” 音講論 (春), pp. 757–760, 3月, 2014年.
- [39] 村瀬 慶和, 千葉 大将, 小野 順貴, 宮部 滋樹, 山田 武志, 牧野 昭二, “伝達関数ゲイン基底 NMF におけるマイク数・マイク配置と目的音強調性能の関係,” 音講論 (秋), pp. 523–526, 9月, 2014年.
- [40] 千葉 大将, 小野 順貴, 宮部 滋樹, 山田 武志, 牧野 昭二, 高橋 祐, “教師なし伝達関数ゲイン基底 NMF による目的音強調における罰則項の特性評価,” 音講論 (秋), pp. 527–530, 9月, 2014年.