

2チャンネルステレオデータからの定位分離

5 P-8

松田 明男, 千種 康民
東京工科大学 情報工学科

1 まえがき

今まで複数音源の定位分離や定位の分布は数々の研究がなされており、マイクロフォン自体のハードウェア的な制限や推定問題であるため数多くの未解決の問題点が残されている。

本研究では、CDやDAT等の既存のステレオ音楽から聴覚的に捕らえる定位分離とその可視化を目指す。また、実現困難な音源分離ではなく、定位の可視化により、聴覚的に捕らえる定位分布をより明瞭に判断することができ、さらに、既存のレベルメータ、グラフィックイコライザに加わる第3のメータ(図1)を構成できる。

2 本研究における音響と定位

音の物理的性質は、強度、周波数というパラメータで表わすことができる。音楽では、さらに音長のパラメータが加わる。ある時刻における音響の定位は、強度、位相によって表現される

2.1 録音環境

録音環境において、演奏楽器とMIDI楽器では定位における相違がある。まず、演奏楽器の場合、複数のマイクで録音すると、マイクと楽器の距離に応じた到達時間と音圧となり、定位を形成する。

一方、MIDI楽器の場合、音源と録音部を直結することが多く、定位操作の機能を有する音源の場合、設定した定位がそのまま伝達される。この定位の多くは音圧の違いのみで行なっており、録音された定位も、また、ミキシング時の定位も同様である。本研究では後者の録音環境を対象としている。

2.2 再生

音楽録音物を聴く手段には2つある。スピーカ出力の場合、LRどちらの出力も左右の耳に音は伝搬される。さらにその伝搬距離が異なるため、位相差が生じる。ヘッドフォンの場合、出力装置が

耳に密着しているため、LRの出力はそれぞれ左右の耳のみに伝搬され、他方の耳へ伝搬されない。本研究では定位分布の可視化において、このヘッドフォンにおける定位を対象としている。

2.3 音源分離と定位分離

音源分離とは入力信号を信号源毎に分離することであり、複数音声からそれぞれの音声を抽出したり、雑音下の中にある特定の音響信号を抽出することである。ここで、入力信号はモノラルデータでも構わない。

一方、定位分離とは複数の入力信号と入力環境から定位毎に分離したり、音源の方向を推定することである。同一方向からの複数音源は混在したままになるが、ある方向に単一の音源しか含まれていなければ、同時に音源分離されることになる。

2.4 定位分離での音圧差

前述のように、ヘッドフォン再生では左右の耳に加わる音圧比によって、定位を制御できるので、既存の音圧で定位制御を実現している。その制御は図2に示す比により実現される。L側音圧を V_l 、R側音圧を V_r とすると、定位によらず同じ大きさの音であるには、 $V_l^2 + V_r^2 = 1^2$ が成立すればよい。L側定位が $\theta = 0$ ならば、

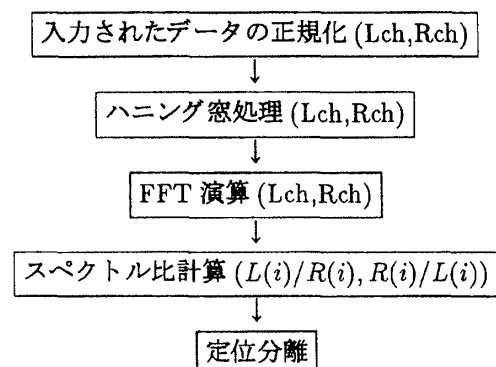
$$V_l = \cos(\theta/2) \quad (0 \leq \theta \leq \pi)$$

$$V_r = \sin(\theta/2) \quad (0 \leq \theta \leq \pi)$$

となる。

2.5 本アルゴリズム

図3に示す本手法は以下の手順で処理をする。



¹Sound Separation System for MIDI Music

²Akio Matsuda and Yasutami Chigusa

³Tokyo Eng. Univ. 1404-1 Katakura, Hachioji, Tokyo 192, Japan, <http://www.teu.ac.jp/diit/>

定位算出の手順は、i番目の左右の振幅スペクトル： $|L(i)|$ 、 $|R(i)|$ 、
 $|L(i)| = \sqrt{Re\{L(i)\}^2 + Im\{L(i)\}^2}$
 $|R(i)| = \sqrt{Re\{R(i)\}^2 + Im\{R(i)\}^2}$
 の比によって定位位置を算出する。i番目の定位を $\theta(i)$ とすると、次式で示すことができる。

$$\theta(i) = \begin{cases} 2 \tan^{-1}(L(i)/R(i)) & (L(i) \leq R(i)) \\ \pi - 2 \tan^{-1}(R(i)/L(i)) & (L(i) > R(i)) \end{cases}$$

この式によってLRに対応するスペクトルから定位 $\theta(i)$ を算出し、その時刻におけるパワースペクトル $(L(i)^2 + R(i)^2)$ として表現される定位に積算される。図3に定位算出方法の概念図を示す。なお、本研究においては、位相スペクトルの値を用いて、その定位の正当性を評価し、正当性の低いものは積算しないようにしている。

3 結果と今後の課題

実験で使用した入力条件は、サンプリング周波数 44.1kHz, 16bit, 2ch ステレオデータ MIDI 音源を用い、直結録音、音源間のスペクトル相関が小さいもの、である。本アルゴリズムにおける定位可視化を、以下音源条件で実験を行ない、検討した。

- 1) 単音源での固定定位、
 - 2) 単音源での動的定位、
 - 3) 複数音源での固定定位、
 - 4) 1つの固定定位音源、1つの動的定位音源。
- 1), 2) は、定位分離の性能評価であり、3), 4) は、代表的な定位分離り、様々な音色、定位に対してシステムを評価した。この結果を図4に示す。

今後の課題は、過去の履歴を利用した算出アルゴリズムの強化、リアルタイム録音に対応である。

参考文献

- (1) 安藤: リアルタイム 3次元音源定位センシング・システム, インタフェース, pp.128-136, Feb. 1995
- (2) 小池, 柳沢: 音圧マイクロホン対による音源方向分離に関する可能性の検証, 信学論 (A), J77-A, 1, pp.1-6(1994)
- (3) 後藤, 村岡: 打楽器音を対象にした音源分離システム, 信学論 (D-II), J77-D-II, 5, pp.901-911(1994)
- (4) 森田: 音響パラメータ探索法による複数話者の音声分離, 信学論 (A), J73-A, 10, pp.1551-1557(1990)

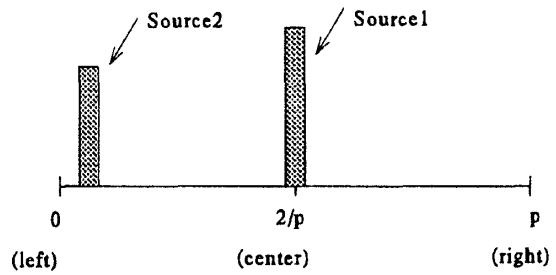


図 1: 定位の可視化 (定位レベルメータ)

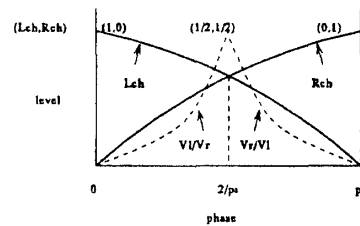


図 2: 左右の位相と音圧比

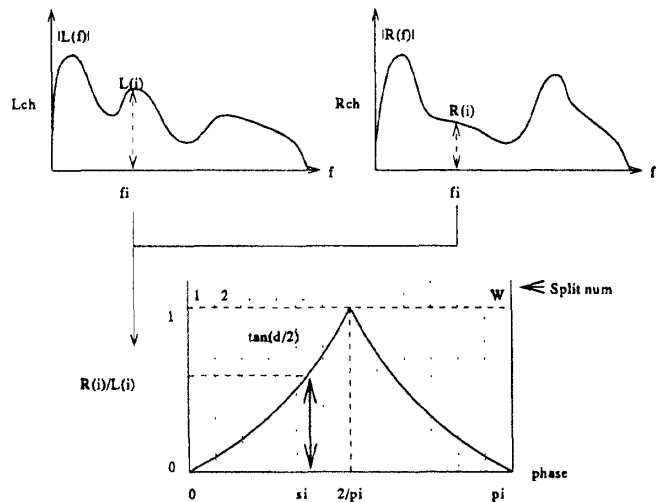


図 3: 定位分離の概念図

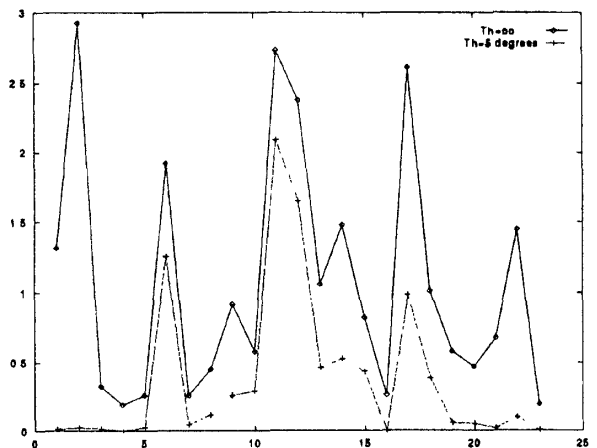


図 4: 定位分布の積算結果