

参加体験型ワークショップにおける 複数の音響信号からの位置関係の抽出

中村 嘉志[†] 友部 博教[†] 須永 剛司^{††} 西村 拓一[†]

[†] 産業技術総合研究所 情報技術研究部門

^{††} 多摩美術大学 情報デザイン学科

E-mail: nmura@ni.aist.go.jp, h-tomobe@aist.go.jp,
sunaga@tamabi.ac.jp, taku@ni.aist.go.jp

あらまし 本稿では、ワークショップ（参加体験型の創造的活動）での活動記録を目的として、記録した複数の音響信号から、参加者のその場における位置履歴を推定する手法について述べる。提案手法は、ワークショップの音声記録を行うだけでなく、位置センサや方向センサを用いなくとも、録音装置と音源装置を用いるだけで参加者の位置履歴を推定することができる。音源装置から発せられた固有音は録音装置で記録され、この固有音を頼りに解析することにより、オブジェクト（ヒトやモノ）のそれぞれの位置関係を抽出する。ユーザは、抽出された位置関係を位置による活動履歴として後で振り返ることができる。本稿では、予備実験を通して提案手法の有効性を評価する。

キーワード 音響による位置推定, 方向推定, ワークショップ活動記録

Recognition of Location Relationships using Audio Signals for Participatory Learning Workshops

Yoshiyuki NAKAMURA[†], Hironori TOMOBE[†], Takeshi SUNAGA^{††},
and Takuichi NISHIMURA[†]

[†] Information Technology Research Institute, National Institute of AIST

^{††} Department of Interaction Design, Tama Art University

E-mail: nmura@ni.aist.go.jp, h-tomobe@aist.go.jp,
sunaga@tamabi.ac.jp, taku@ni.aist.go.jp

Abstract This paper proposes an activity capture method of attendees and facilitators for indoor interactive workshops, which are events designed for participatory learning and creative endeavors in group. Without any special location sensors or direction sensors, a simple acoustic recorder and player for every user and artifact in the workshop enables the method to estimate the user location history as well as recording the audio scenery. Each audio signal captured by a recorder is analyzed and identified as a specific sound emitted from a corresponding audio player. The locations and orientations of all users are estimated by collecting all the information in the vicinity of each attendee. Users can re-experience the workshop audially and visually using a map of the workshop room and attendees' locations and orientations. Our preliminary experimental results demonstrate the feasibility of our estimation method.

Key words Acoustic location estimation, orientation estimation, workshop activity log

1 はじめに

1960年代以降、世界中の様々な分野において、ワークショップと呼ばれる参加体験型の学習や創造的活動が広がりを見せている [1]。このようなワークショップは、合意形成や創造表現のみならず、ブレインストーミングや参加者同士の体験共有、コミュニケーションを活性化する効果があると考えられる。我々はこのようなワークショップでは、参加者や主催者にとってリフレクション（振り返り）が重要であると考えている。

リフレクションとは、参加者や主催者がワークショップの活動記録を閲覧し、活動に解釈や感想を付与し、自己や他者が付与した解釈や感想を閲覧することで、活動を振り返り思考することと本論文では定義する。そのための電子的なツールとして、我々はワークショップリフレクターと呼ぶシステムを提案している [2]。これは、ワークショップで記録した写真や映像記録などの情報を整理・加工して、参加者および主催者に時系列、またはKJ法 [3] に似たカード型情報としてそれらを提示するツールである。

本稿では、このワークショップリフレクターを用いて、特に主催者がワークショップそれ自体の社会文化的な分析行為を実施することに焦点を当て、そのために必要な活動記録を行う音声を利用した記録手法を提案する。すなわち、参加者同士がワークショップ中でどのようなインタラクションを行い、それらによってどのように学習行為や創造的活動が進んだかをエスノグラフィによる会話分析手法などにより分析可能とさせるシステムを提案することが本稿の目的である。

社会文化的な分析においては、彼らの発話によるインタラクションのパターンが特に重要とされる [4]。そこで本稿では、複数地点で同時に記録した複数の音声情報から参加者同士の位置関係、つまり、どの参加者がどのような近さと向きで会話をしたかを抽出する手法について提案し、予備実験を通じて提案手法の特性を述べる。

2 音響信号を用いた位置関係抽出手法の提案

2.1 ワorkshopにおけるシナリオ

本稿で対象とするワークショップは、室内で実施される参加者数十人に規模の参加体験型創造活動である。これは、ファシリテータと呼ばれるワークショッ



図1 ワorkshopにおける作品の例：
グループまたは個人単位でテーマにした
がった作品を製作する

Fig. 1 An example of workshop outputs.

プの設計および司会者が予め参加者に対してテーマを設け、それに従って参加者がグループもしくは個人で作品を仕上げる過程を通して、学習やコミュニケーションの活性化を図るものである。例えば、予め用意してある多数の写真の中から数枚を使って、図1のように川柳のような詩にまとめるものがある。

図2は、我々が開発しているワークショップリフレクター [2] を用いた振り返りまでの過程を示したものである。このようにワークショップでは、通常、実施中に様々な記録デバイスを使用して活動の記録を行い、この記録から主催者は最終的に社会文化的な分析を行う。

こうした背景の下、我々はそのベースとなる音声録音に着目した位置関係取得手法を提案した [5]。本稿では、これを発展・具体化させ、位置センサーや方位センサーなどの音響デバイス以外のセンサー機器を用いなくても参加者相互の位置関係と会話によるインタラクション記録を同時に行うシステムを提案する。

2.2 音響信号を用いた位置関係抽出手法

提案手法では、図3に示すように、参加者同士や壁や机などの基準点との間の二者間の相対位置関係を取得する。相対位置関係とは、本稿では、相対的な距離および角度の関係と定義する。取得したこれらの位置関係は、これまでに我々が提案しているトポロジー推定 [6] を用いて全体の位置関係、すなわち絶対的な位置推定へと繋がるものとする。ここでは、まず、基礎となる相対位置関係をどのように取得す

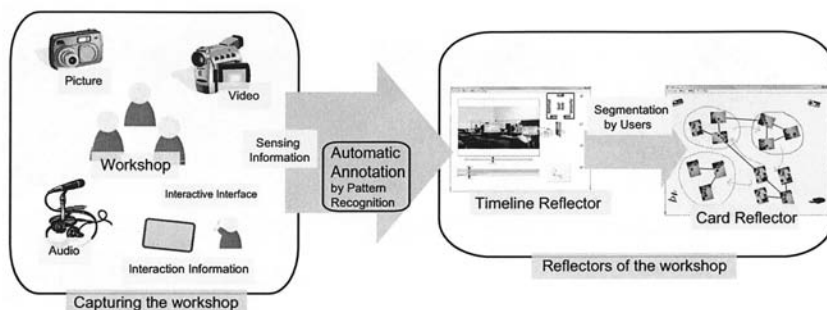


図2 ワークショップ・リフレクターの処理の流れ
Fig. 2 Flow of workshop reflector.

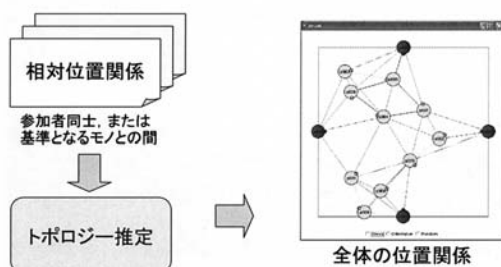


図3 提案手法のフロー図
Fig. 3 System flow of proposed method.

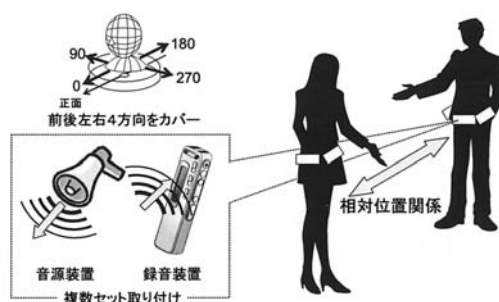


図4 提案手法のデバイス構成
Fig. 4 System structure.

るかについて述べる。

位置関係を取得するためには、図4に示すように、ワークショップの参加者がそれぞれICレコーダに代表される録音機器を装備し、実施中は自身の発話、そして周囲の音を絶えず録音することを基本とする。このとき、音響マイクをそれぞれ異なる方向に向けておく。音響マイクには指向性があるので、こうすることでマイクを向けられている方向の音が強く録音されることになる。図では、前後左右の4方向に対してマイクが向けられていることを示している。

一方、図4に示したように、マイクと一対になるように音源装置、つまりスピーカを設置する。このスピーカからは、それぞれの固有音が定期的に出力されるようにしておく。固有音は、フィルタ処理しやすい正弦波や、任意の情報をデコード可能な音のQRコード[7]などを出力する。参加者自身の声(声紋)の違いを固有音として音源装置を省略することも考えられるが、解析処理の簡単化のためと、前方以外の方向への出力および位置関係抽出に対応するため、ここでは積極的に音源装置を用いることとする。

音は音速で伝わるので、固有音の録音時刻は、一対としたマイクでの録音時刻とそれ以外のマイクでの録音時刻—例えば図4で言う向き合っている相手—の間では時差が生じる。この固有音の録音時差が二者間の相対距離に(スピーカと一対となるマイクとの距離は取得したい相対距離に比べて無視できるものとする)、マイクの録音指向性による固有音の録音強度が相対角度となる。こうして録音された複数の録音データを一つ一つ解析することにより、ある参加者から見ていつ、どの方向のどの位置に他の参加者(もしくは壁や机などの基準点)が存在するか、それぞれの位置関係を取得する。

2.3 トポロジー推定の概要

2.2節の手法により得られた複数の位置関係は、トポロジー推定[6]により全体の位置関係へとマッピングする。ここでは、トポロジー推定についてその概略を述べる。

トポロジー推定は、位置センサ(GPSなど)や向きセンサ(電子コンパスなど)を用いなくとも、対象

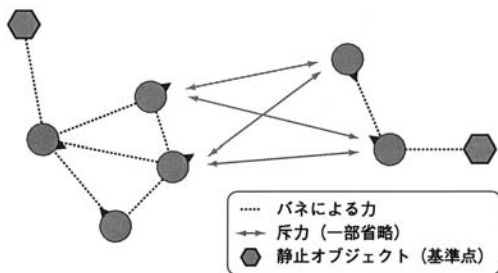


図5 トポロジー推定の拡張バネ力学モデル
Fig. 5 Extended spring dynamical model.

とするオブジェクト（ヒトやモノ）全体の位置および向きを推定できる特徴を有する。また、相対位置関係の数が多ければ多いほど推定精度は向上するため、比較的混んだ空間で利用することが可能である。

トポロジー推定の考え方は以下の通りである。取得した相対位置関係のうち、推定対象空間内において位置および方向が既知のオブジェクトがいくつか分かれば、その位置を基準として他のオブジェクトの位置を相対的かつ再帰的に規定することができる。そのための推定のアルゴリズムとして、トポロジー推定は、図5に示すようなバネと斥力を用いた力学モデルを用いている。取得された位置関係は、バネに置き換えられてバネ力で結び付く。加えて、回転方向に対してもバネ力は働いて、方向に対する推定に寄与する。一方で、位置関係が取得できないオブジェクト同士には斥力が働く。これは、関係を持たないオブジェクトは遠くにいるだろうという仮定による。

トポロジー推定では、バネ力と斥力の総和が局所解に落ち着く場合を最も矛盾が少ないと定義し、これによって二次元平面上でのオブジェクト全体の位置および方向を規定する。

2.4 録音データの時間同期

録音した複数のデータ間では、サンプリング周波数のバラツキによる時間のズレが生じる。長時間録音した場合、これがドリフト誤差として顕在化し、相対位置関係を正しく抽出できなくなる。実際、ローランド社製 EDIROL R-09 ICレコーダを用いて実測したところ、個体差により2時間あたり最大で約14ミリ秒の誤差が生じた。このため、位置関係抽出処理時には録音データ間で時間におけるキャリブレーションを絶えず行なう必要がある。

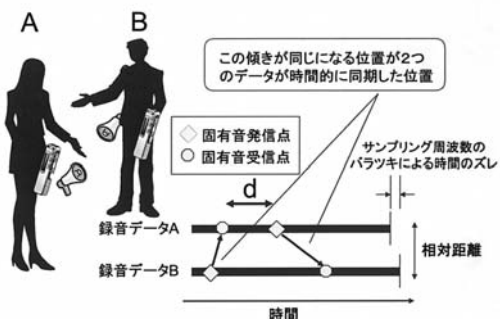


図6 録音に関する時間誤差の補正
Fig. 6 Time calibration for recorded data.

本研究におけるキャリブレーションは、固有音におけるToF(Time of Flight)法を用いた録音データ間の時間相関を検出することで行う。いま、図6に示すように、二者間で固有音の信号をやりとりする場合を考える。参加者、つまり人間の移動は音速と比較して十分に遅いので、固有音をやりとりした時間差（図中では d で示した差）が短ければ、移動誤差は無視できて二者は停滞していると仮定することができる。したがって、 $A \rightarrow B$ 、 $B \rightarrow A$ の固有音の伝達時間は同じとなる。この性質を利用して録音データ間の時間相関を検出することで補正を行う。

3 位置関係取得実験と評価

実際のICレコーダとスピーカを用いて、二者間の位置関係（相対距離と相対角度）がどの程度取れるか実験して評価した。本節では、その実験の概要と結果について述べる。

3.1 信号音とバンドパスフィルタ

信号の発信者および発信方向を特定する固有音として、実験では、16kHz、約2ミリ秒の長さの正弦波を用いることとした。これは、人間の声の中心周波数からは十分に離れているため、会話と信号の分離が容易であるためである。また、この信号音に対応する、表1に示す特性と係数を持った、カイザー窓を用いたFIR（直線位相）型の音響フィルタを施して、録音データの中から信号音を抽出することとする。

3.2 実験方法

実験では、二者間の位置関係、つまり相対距離と相対角度を提案手法により求め、評価を行う。以下では、それぞれについて使用機器と実験方法を示す。

表1 音響フィルタの特性
Table 1 Specifications of acoustic filter.

カットオフ周波数 (fc1)	15.5kHz
カットオフ周波数 (fc2)	16.5kHz
サンプリング周波数 (fs)	44.1kHz
タップ数 (n)	50

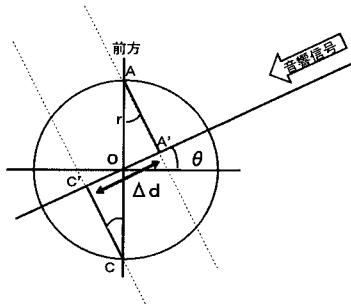


図7 相対角度の算出方法
Fig.7 Method of relative orientation estimation.

信号音の音源装置として、Apple社製 iPod shuffle ミュージックプレーヤとビクター社製 HP-AL202 耳掛け型スピーカを使用する。また、録音装置として、ローランド社製 EDIROL R-09 IC レコーダにパナソニック社製 RP-VC151 マイクロフォンを装備したものをを用いる。

相対距離を求める実験では、録音装置を2台使用する。2台のうち1台のマイクを音源装置の傍ら、3cm離れた場所に固定し(これをAとする)、もう1台の位置を変化させる(これをBとする)。AおよびBはそれぞれ異なる録音装置であるため、音源装置から発せられた信号音Sは、別々の録音データファイルとしてそれぞれ独立に記録される。AとBの録音開始時刻を同期させておけば、AおよびBにおける信号音Sの記録時間差 Δt が求まる。ToF法により、この Δt を相対距離へ換算する。

相対角度を求める実験では、1台のICレコーダにLチャンネル、Rチャンネルそれぞれ別々のマイクを接続し、これらが直立した円筒形を上から見た場合に前方および後方にそれぞれ向いているものとする。これは、ベルトをまとった人間の臍の部分と背中部分にマイクがそれぞれ設置されているものを仮定している。この円筒形の円平面上で、音源装置の位置をずらして信号音Sの入射角を変化させる。入射角に応じてLチャンネルおよびRチャンネルに記録される信号音Sには時間差 Δt が生じるので、これを図

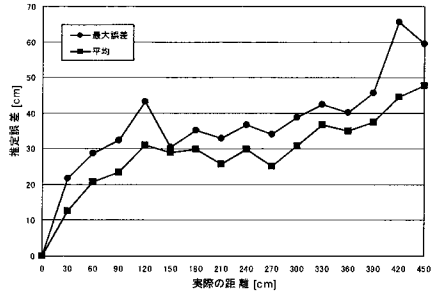


図8 相対距離推定の結果
Fig.8 Result of relative distance estimation.

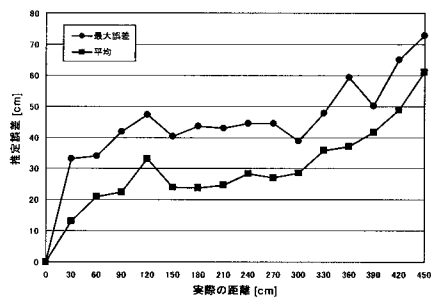


図9 相対距離推定(会話ノイズ含む)の結果
Fig.9 Result of relative distance estimation with noise.

7で示す幾何学的近似により相対角度を求める。図中のAとCはマイクの設置位置であり、音源装置の位置がマイクの設置円平面の半径rよりも十分に大きく、かつ、信号音Sが平面波として伝わると仮定すると、 $\Delta d = 2r \sin \theta$ で近似できる。音速を考慮した Δd と Δt の相関から相対角度 θ を算出する。

3.3 結果および考察

まず、相対距離の測定結果を、図8および図9にそれぞれ示す。X軸は、スピーカとマイクを正対させた状態での実際の相対距離であり、0.3mごとに0mから6mまで測定したものである。Y軸は、3.2節で述べたToF法にしたがって算出した相対距離と実際の相対距離とのズレを推定誤差として求めたものである。両結果共に10回測定の実験結果である。前者の図8は、ホワイトノイズ約40dBAの静かな空間で計測した結果で、後者の図9は、ワークショップでの参加者の会話を考慮して、ニュース朗読を騒音レベル80dBAで流した状態で測定したものである。

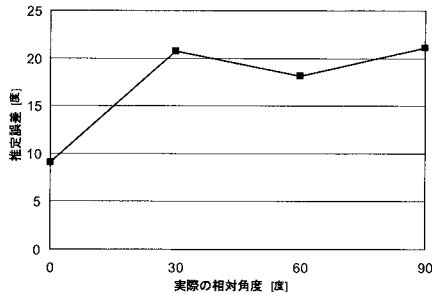


図 10 相対角度推定の結果

Fig. 10 Result of relative orientation estimation.

相対距離の測定結果より、会話がある場合、ない場合共に、距離 4m まで誤差 0.5m 以内で相対距離を測定可能であることが分かった。これは、参加者を中心に考えた場合、自身に関係した 4m 以内のヒトやモノが把握できることを示しており、会話相手を中心に行う社会文化的な分析に適用できると期待できる。また、これまでのトポロジー推定 [6] では、5m 以内の範囲に相手が居るか居ないかの情報しか考慮していないため、0.5m の誤差で相対距離が測定できることは、トポロジー推定の精度向上にも繋がるものと期待できる。なお、相対距離が 4.5m を越えると S/N 比が悪くなり、本稿の手法では推定することができないため別の工夫が必要であることも分かった。

次に、相対角度の測定結果を図 10 に示す。実験では実際の相対角度 θ を 0 度から 90 度まで 30 度ごとに変化させて角度分解能を測定した。信号音 S の音源装置を距離を 1m 離れた場所に設置した。3.2 節で述べた半径 r は、人間の胴回りを考慮して 0.15m とした。結果は、それぞれの角度で 10 回ずつ計測を行った平均である。なお、相対距離の測定同様に、実際の会話を考慮したニュース朗読雑音がある場合とない場合で計測したが、特に優位な違いは認められなかったため、ここではホワイトノイズのみの結果を示す。

相対角度の測定結果より、誤差約 20 度以内で角度推定できることが分かった。人間は興味のあるモノの方向を向く性質があるが、そういった情景を取得できることが分かった。したがって、社会文化的な分析に適用できると考えられる。また、これまでのトポロジー推定の実験では、相対角度は 90 度単位でしか捉えていなかったため、本提案手法を用いるこ

とにより、相対距離と共にトポロジー推定の精度向上が期待できる。

4 おわりに

本稿では、ワークショップの活動記録における音声記録に着目し、位置センサーや方位センサーなどの音響デバイス以外のセンサー機器を用いなくても参加者相互の位置関係と会話によるインタラクション記録を同時に行う手法を提案した。市販されている実際の機器を用いて実験を行って、提案手法により音声記録の中から位置関係を抽出できることを示した。

今後は、音源装置を用いなくても、録音された複数のデータの中から似通った音を見つけて相対位置関係を抽出する情景分析 (Scene Analysis) を行ってみたいと考えている。また、推定された位置情報と複数の音声情報を用いて、位置に基づいた音のモーフィングによる再生技術について検討したいと考えている。

謝辞 本研究の一部は、JST CREST 「情報デザインによる市民芸術創出プラットフォームの構築」の助成を受けて行った。また、実験を遂行してくれた早稲田大学理工学部 4 年の一柳智史氏に謝意を表す。

参考文献

- [1] 中野民夫：ワークショップ - 新しい学びと創造の場、岩波書店、ISBN:4-00-430710-4 (2001)。
- [2] Tomobe, H., Nakamura, Y., Numa, K., Nishimura, T. and Sunaga, T.: Workshop Reflector: an Interface for Reflecting Processes of Human Activities on Workshops, *In Proc. of ACM CHI2008(Work-in-Progress)*, to be appeared (2008)。
- [3] 川喜田二郎：発想法 - 創造性開発のために、中央公論社、ISBN:4-12-100136-2 (1967)。
- [4] Boden, D.: *Talk and Social Structure: Studies in Ethnomethodology and Conversation Analysis*, Univ of California Press, ISBN:0-52-007506-4 (1991)。
- [5] Nishimura, T., Nakamura, Y., Tomobe, H., Kurata, T., Okuma, T. and Matsuo, Y.: Location Estimation using Auditory Signal Emitted and Received by All Objects, *In Proc. of International Conference on Networked Sensing Systems(INSS2007)*, p. 295 (2007)。
- [6] 中村嘉志, 並松祐子, 宮崎伸夫, 松尾 豊, 西村拓一: 複数の赤外線タグを用いた相対位置関係からのトポロジカルな位置および方向の推定, 情報処理学会論文誌, Vol. 48, No. 3, pp. 1349-1360 (2007)。
- [7] NTT DoCoMo 報道発表資料:音響 OFDM, <http://www.nttdocomo.co.jp/info/news.release/page/20060413.html>。