

# 複数人の指差しを用いた車内会話の場所への紐付け

坂本 匡<sup>†1</sup> 松村 耕平<sup>†1</sup> 野間 春生<sup>†1</sup>

**概要:** 本研究では自動車内での環境を想定して、指差し行動を認識し、その指示対象を自動識別することで、その地点と車内での会話との紐付けを行う手法を提案する。複数人に対して小型のモーションキャプチャ装置を用いて位置姿勢の計測を行い、手と頭の位置から直線の式を求めて、複数の直線から交点を求めることで指差し対象を同定する。本稿では、提案手法による識別が可能かを検証する実験を行ったのでその結果について考察する。

## Embedding In-car Conversations into Places using Finger-pointing

TADASHI SAKAMOTO<sup>†1</sup> KOHEI MATSUMURA<sup>†1</sup>  
HARUO NOMA<sup>†1</sup>

**Abstract:** In this study, assuming the environment in a car, we propose a method of embedding in-car conversations into places by using recognition of finger pointing. We use a motion capture device towards people and measure their positions and attitudes, and we identify the object of their finger pointing by using cross point of their finger pointing. In this paper, we discuss the result of the experiment that validates whether the pointed object can be recognized by our method or not.

### 1. はじめに

我々がある対象について会話をするとき、指示語と共にしばしば対象を指差すことがある。我々が、傍観者として会話を眺めるとき、そのような指差しを目にすれば、指差された対象について会話をしていると理解することができる。このことを逆に捉えれば、指差しという行為を抽出することが出来れば、ある対象についての会話だけを取り出すことが出来る。例えば、友人と歩きながら「そのカフェの雰囲気が素敵」と言いながら指差しを行っていればその指差した先のカフェの口コミとなるし、ドライブ中に「あその交差点はよく人が飛び出してくるから危険」という話の途中に指差しがあれば、それは危険な地点の情報となる。また、同じ場所であっても、例えば、「この道は春になると桜が綺麗」や、「夜になるとイルミネーションが美しい」など、季節あるいは時間によって会話が変化する。このように、会話する内容は場所、季節、時間、あるいは会話者の関係性などによっても変化し、シチュエーションに強く結びついている[1]。

会話の内容を、指差しという動作によって対象の地点と自動的に結びつけることが出来れば、時間などの状況と場所に結びついた情報の一つとして捉えることが出来る。そして、これらの情報を人々に流通させることで、その場所に関する知識の共有や、共有によって得た知識に上乗せする形での知識の創出、あるいは知識を得たことによる行動の変容などの効果が期待できる。

本研究では、そのことに着目し、目的とする。本論文では、特に車内での会話に着目し、複数人の指差し行為による指差し対象の特定手法について述べる。

### 2. 関連研究

本論文で対象とする、人のジェスチャに着目した研究としては Mistry と Maes によるカメラとプロジェクタを用いたユーザインタフェースに関する研究などが有名である[4, 5]。特に、自動車に関連した研究に関しては、Ohn-Bar らの車内での情報端末へのアクセスに手のジェスチャを用いた研究[6]や、Riener らの運転者が車内でのジェスチャを行う空間に着目し、車内における有効なモーションキャプチャの方法を提案した研究[7]などがある。Rümelin らの研究[8]では、本稿と同じように車内における指差しについて検証を行っている。彼の手法では、Microsoft Kinect を用いて計測した手と頭部の位置から指差しのベクトルを計算し、そのベクトルと車の前方 2m の面との交点を指差しの対象としていた。しかしながら、指差しの対象が前方だけにあるわけではなく、また、同じ方向に複数のランドマークがある場合、そのどれを指しているのかをこの手法では識別することができない。これらのような課題を改善するため、我々は次章の手法を提案する。

<sup>†1</sup> 立命館大学  
Ritsumeikan University

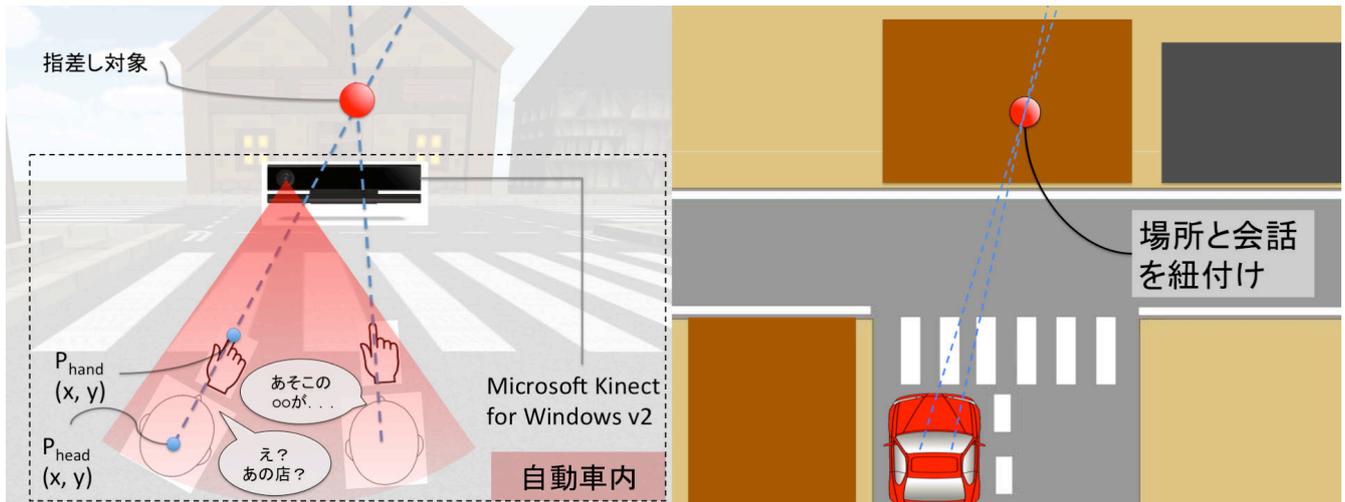


図1 提案手法の概要

### 3. 提案手法

我々は、人々が会話の際にその対象を指さすことに着目する。図1左は提案手法の概要を図示したものである。複数人による指差しによって、会話を紐付ける場所の同定をする。

我々の提案する手法では小型のモーションキャプチャ装置を用いて身体動作の計測を行い、指差しの認識を行う。指差し認識のプロセスは以下である。(1) 指差しの判断：モーションキャプチャ装置から手の位置、頭の位置を得て、手の位置が肩の高さにある時に指差しの検出を開始する。(2) 図1左のように搭乗者それぞれの頭の位置と手の位置を入力として、それらを結ぶ直線を求める。

(3) (2) で求められた2つの直線を求める。この交点を搭乗者が指さした対象とする。従来の手法と比較すると、深度の情報も取れるほか、複数人の情報を用いているためより高い精度が期待できる。

また、これらの操作によって得られた座標の単位をメートルから経緯度のオフセット量に変換し、現在の車輛の座標・方角に対して足し合わせることで、地図上における指差し対象の経度と緯度を決定する(図1右)。このように決定された緯度経度の地点と、指差しがされていた時の会話を紐付けし、活用する。

### 4. 評価実験

#### 4.1 実験概要

この手法による識別精度を検証するため、我々は次のような実験を行った。被験者は20歳から22歳の大学生8名である。被験者には二名一組のペアを作り、並んで立って特定の位置にある対象を指さすことを依頼した。指示対象には視認性を考慮して硬式テニスのボールを採用した。指示対象は図2に示すように、被験者から3m・5m・10mの地点にそれぞれ11点、11点、9点を放射状に配置した。3mおよび5mの点については前方100°に10°刻み、10m

の点については前方40°に5°刻みの位置に配置した。配置した計31点につきそれぞれ4回、事前に決定したランダムな順で2秒間の指差しをさせた。各ペアで左右の立ち位置を変更して、2セットを行った。このときの被験者2名のボディモーションをMicrosoft Kinect v2[a]を用いて記録し、頭の位置と手の位置について平均を取った後に前述の直線の式を求めた。2名の被験者から求められた直線の交点が、その試行における指差し対象の位置となる。

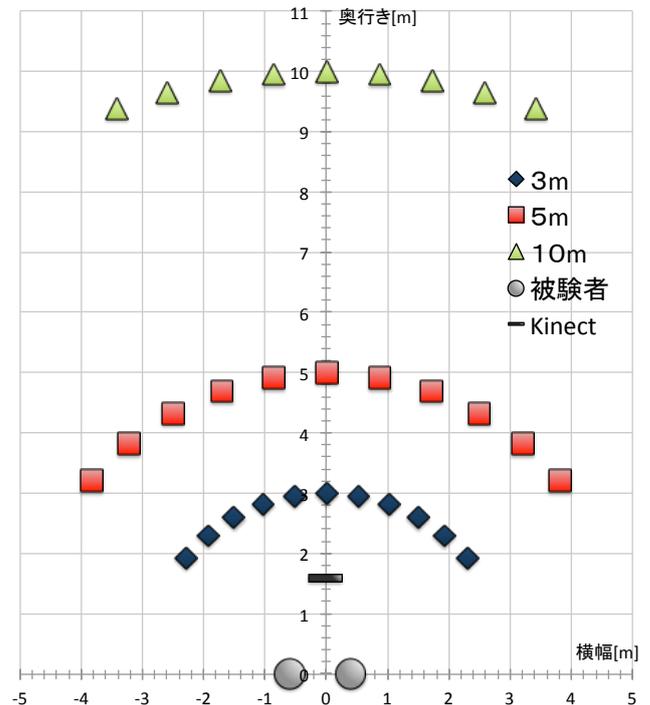


図2 指差し対象の配置図

#### 4.2 実験結果

4ペア8セットの実験から、各地点について32回の計測データ、総数992試行のデータが計測できた。各試行によって得られた指差した交点の位置から上下5%ずつを除外

a) Microsoft Kinect for Windows v2  
<http://www.microsoft.com/en-us/kinectforwindows/default.aspx>

し、指定された点と指差した交点との距離誤差について分析を行った。まず、得られた距離の誤差の平均、分散について表1に示す。中央値についても算出したが、平均値と差は無かったため、表からは除外した。

表1 各距離の点における距離の誤差

		左側の点	中央の点	右側の点	全体
平均	3m	0.44m	0.39m	0.38m	0.40m
	5m	1.09m	1.02m	1.02m	1.05m
	10m	2.46m	2.65m	2.69m	2.59m
分散	3m	0.03	0.02	0.04	0.03
	5m	0.26	0.25	0.32	0.28
	10m	1.41	1.25	1.42	1.36

また、各距離の点について、指差すように指定された指示対象の位置と指差された交点をプロットしたものを図3、図4、図5に示す。大きな点が指差すように指示された指差し対象であり、小さな点群が実際に指差された点である。同じ形・色をした点は同一の指示対象を指差したものである。なお、視認性向上のため、点を半分に間引いて示した。

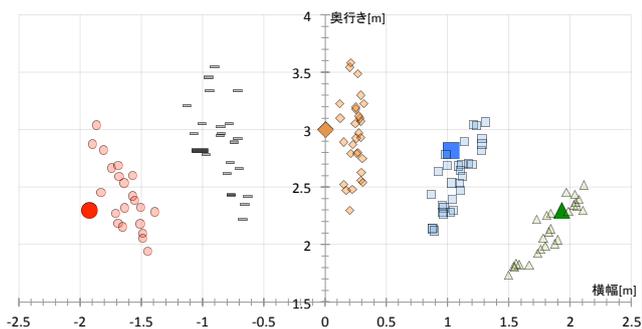


図3 3m地点における指定点と指差された点

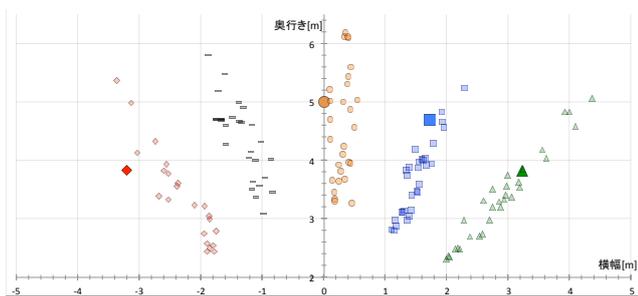


図4 5m地点における指定点と指差された点

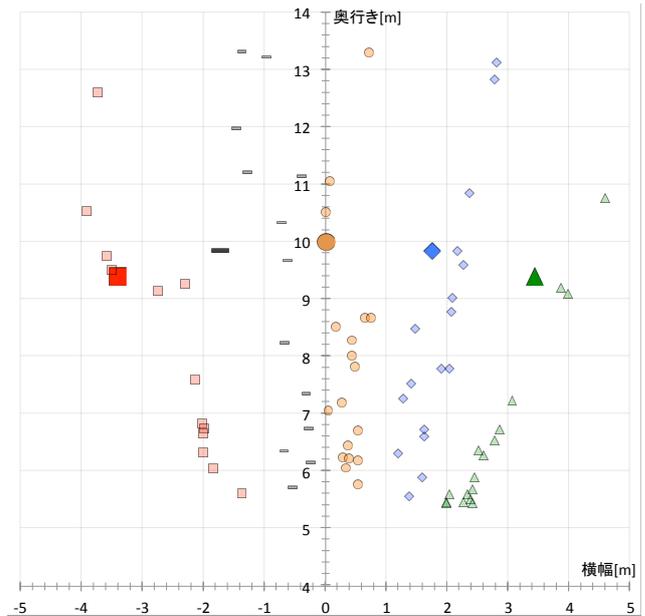


図5 10m地点における指定点と指差された点

表を見て分かる通り、左側・中央・右側で大きな差は無い。距離ごとの差について分析する。指差しを行う人から遠くに離れた点ほど、誤差は大きくなっている。また、プロットした図を見ると、全体的に、特に左側の点において、指差し対象となる点よりも右側を指していることがわかる。

#### 4.3 考察

前述の通り、距離が遠くなるほど誤差が大きくなる。3mのとき0.40m、5mのとき1.05m、10mのとき2.59mの誤差は、一見するとほぼ直線状に並んでおり、一次関数的に増加していくように見えるが、一次式で補間すると0mのときの誤差が負になるため、一次関数的では有り得ない。0m地点における誤差を仮に0mであるとすると、この誤差は距離の二乗に比例していると言え、一般的なビルの底面を20m×20mの正方形であるとした場合、約24m先のビルの中心を指差しても、平均的な誤差であれば、指差した点はビルの中に収まると言える。このことから、車から25mほどの点までに関しては、この手法が有効であり、指差し対象の認識は可能であると考えられる。

また、図から、指差しを行う人を中心とした放射状に点の誤差が分布していることが分かる。さらに、指差した時の交点はあまり左右には誤差がなく、前後に対して大きく誤差があることが見て取れる。特に被験者に対して正面の点ではこの誤差が顕著である。この前後の誤差を生み出す要因として考えられるのがKinectの設置方法である。本実験においては、指差しを行う人の全身が映るように指差しを行う人たちの前方1.6m、肩の高さにKinectを設置した。しかし、この位置からモーションキャプチャを行うことにより、真正面を指した場合は突き出された腕によって手首や肘、肩などが隠されてしまい、手の位置が安定してキャプチャできていなかった。手の位置が前後に微妙に変

化し続けることで引き起こされる直線の傾きの微妙な変化は、近距離の点を指し示す場合においてはあまり大きな距離の誤差となって現れることはないが、遠距離の点を指し示す場合においては大きな問題となる。遠距離を指し示す際は複数の直線の傾きが平行に近いので、左右に少しでもずれると、少しの傾きの変化が大きな距離の誤差の要因になる。

この問題を解決する方法として考えられるのが、搭乗者の上からモーションキャプチャを行うことである。もちろん、指差しを行う人々の距離を離して指差しを行わせれば解決できる問題であり、そちらの方が容易でもある。しかしながら、車内という狭い空間においてはそれほどの距離を取ることができないわけではない。従って上から遮られないように撮るのが良いと考えた。ほとんどの車には屋根があるため、設置は容易であると考えた。Kinectの深度方向の精度が悪くとも、今回の手法では指差しを行う人に対する上下方向の位置座標は無視しているため、問題にはならないと考える。

## 5. おわりに

本稿では、複数人の指差しによる指示対象識別手法について述べ、想定する環境における運用可能性を示した。今後はシミュレータを用いた、より実際に近い環境での実験や、実際の車内での実験によって本手法の運用可能性を検証したいと考えている。また、同時に指差さずとも、誰かの指差しによって他者の指差しが誘発され、時間差で指差し行動を行う場合や、車が動いている間に一人が継続して指差しを行う場合が考えられるため、GPSを用いて車輪の変位を計測するなどし、時間差の指差しにも対応できるよう改善を加えたい。また、車内において、Kinectなどの赤外光を用いる計測機器は高い精度を期待できず、ジャイロセンサなどによる計測は高い精度があるものの、装着性が大きな課題となる。そのため、動きの計測方法についても考える必要がある。

さらに、指差し行動による対象の特定と会話の紐付けには、本論文で述べた以外にも応用可能性がある。例えば、Web上の情報と場所の情報は、既にGoogle Maps[b]やYahoo!地図[c]などで紐付けされている。本手法で会話と対象の場所の情報を紐付けることができれば、Webと会話も紐付いた、新たな形のメディアが成立する。今後は、それらの応用系についても検討をしていきたい。

## 参考文献

- 1) Adato, A.: "Occasionality" as a constituent feature of the known-in-common character of topics, *Human Studies*, Vol.3, No.1, pp. 47-64 (1980).
- 2) Kohei Matsumura, Yasuyuki Sumi: What Are You Talking About While Driving? : An Analysis of In-car Conversations Aimed at Conversation Sharing, *Proceedings of the 6th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 1-4 (2014)
- 3) Kohei Matsumura, Yasuyuki Sumi: CarCast: a framework for situated in-car conversation sharing, *Proceedings of the 5th Augmented Human International Conference*, No.17 (2014)
- 4) Mistry Pranav, Pattie Maes, and Liyan Chang. : WUW - wear Ur world, *Proceedings of the 27th international conference extended abstracts on Human factors in computing systems - CHI EA '09*. Boston, MA, USA, 4111. (2009)
- 5) Pranav Mistry and Pattie Maes. 2009. SixthSense: a wearable gestural interface. In *ACM SIGGRAPH ASIA 2009 Sketches (SIGGRAPH ASIA '09)*. ACM, New York, NY, USA, , Article 11, 1 pages. DOI=10.1145/1667146.1667160
- 6) Eshed Ohn-Bar, Cuong Tran, and Mahan Trivedi. : Hand Gesture-based Visual User Interface for Infotainment, *Proceedings of the 4th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp.111-115 (2012)
- 7) A. Riener, A. Frescha, F. Bachmair, P. Hagmuller, A. Lemme, D. Muttenthaler, D. Puhlinger, H. Rongner, A. Tappe, and F. Weger. : Standardization of the In-Car Gesture Interaction Space, *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 14-21 (2013)
- 8) Sonja Rümelin, Chadly Marouane, and Andreas Butz. : Free-hand pointing for identification and interaction with distant objects, *Proceedings of the 5th International Conference on Automotive User Interfaces and Interactive Vehicular Applications*, pp. 40-47 (2013)

---

b) Google Maps  
<https://www.google.com/maps>  
c) Yahoo!地図  
<http://map.yahoo.co.jp/>