

指差し動作認識及び風景認識による 知的なカーナビゲーションシステムに関する研究 A Study of an Intelligent Car-navigation system with Recognition of Hand Signal and Road Scene

澤野 弘明[†] 白根 愛佳^{††} 水野 慎士[†]

Hiroaki SAWANO[†], Aika SHIRANE^{††}, Shinji MIZUNO[†]

We propose a car navigation system considering the passenger's knowledge of route guidance. Our approach is to recognize a hand signal of the passenger and a road scene by a CV (computer vision) technique. First, an inside camera on dashboard takes a video of the passenger. The hand signal is recognized by a CV technique, and a fingertip coordinate is calculated. The fingertip coordinate is transferred to the directed coordinate in the road scene taken by the outside camera on a front grille. The directed coordinate is utilized to calculate the directed region including a target object. A distance is calculated from the object, and the route is recalculated as a navigation system. In this paper, we experiment the fingertip coordinate recognition and the directed region calculation. Depth value from an inside camera to the passenger is used with Kinect. A road scene CG video is used for an ideal experiment. We ask seven examinees to indicate sign boards after camera calibration in advance. The successful ratio of the total is 47% because the view point of the user is not considered. The issue lists for practical system are concluded for future works.

1. はじめに

カーナビゲーションシステム(以下、カーナビ)の機能は、インターネットサービスの利用や、スマートフォンとの連携などにより日々進化しており、普及率も大幅に増加している。そこで様々なユーザに対応するために、円滑に移動するための経路探索、使いやすいユーザインタフェース [1] が求められている。例えば、ユーザを目的地に円滑に到着させるために、走行中に道路交通情報通信システム (VICS: Vehicle Information and Communication System) やプローブ情報が利用されており、渋滞回避を考慮した推奨経路がユーザに提示される。しかし、カーナビによって提示された推奨経路よりも、ユーザの経験による抜け道の方が早く到着する場合があることや、地図の劣化により正確な経路を提示できない場合がある。運転手の経験を利用する場合は運転手の意志により経路を変更することが可能だが、助手席ユーザの経験を利用する場合は運転手に目的の交差点を伝える必要がある。しかし、現在のカーナビの欠点は、助手席ユーザとの会話における「そこ」、「ここ」、「あっち方向」、「こう行ってこう行く」など、人間対人間同士では自然に行なっている指示方法が使いにくいことである。これは助手席によく道を知る人がいる場合ですら、時に車を

停めてめんどうな操作を行わなければならないことがある。

そこでこの課題を解決するために、システムが助手席ユーザの意図を認識し、推奨経路を動的に変更する手法に着目する。これまでの筆者らの報告では、助手席ユーザが指示した対象物体を CV (Computer Vision) 技術により認識し、認識結果に基づいてカーナビの推奨経路を変更するシステム [2, 3] の基礎検討を行なってきた。ユーザ動作をカーナビ等の操作に活用する研究例は他にも報告されている。OKI セミコンダクタ株式会社や Audi は赤外線 LED カメラを用いて手の形状と動きを認識し、カーナビ操作に特化して機器を操作する製品 [4, 5] を提案している。しかし、これらの製品では、単純な手の形状に依存するため、操作性が限定されており、ジェスチャーの習得が必要である。また、Fang らは色と動作を考慮した手の動作認識法 [7] を提案している。色情報を利用した場合、指と顔の色相が近いため、誤検出が多くなると考えられる。

そこで我々はカメラからの距離(深度)情報を用いた指差し動作の認識法及び認識結果に基づいた風景中の指示領域推定法を提案する。また、試作システムは指差し動作認識に基づいた、対象物体周辺の領域を抽出する簡易的な機能のみ実装し、案内機能は実装されていないため、本論文では、本システムの実装時に考慮すべき点と問題点を列挙し、本システムを実装するための指針を示す。

[†] 愛知工業大学情報科学部情報科学科 (豊田市)

^{††} 株式会社瓢屋 (名古屋市)

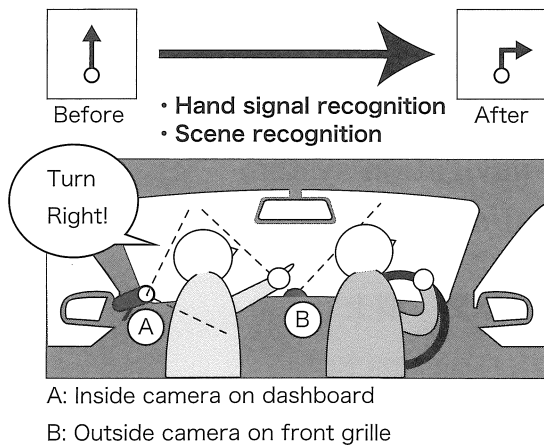


Fig. 1. An application to route guidance by recognition of hand signal

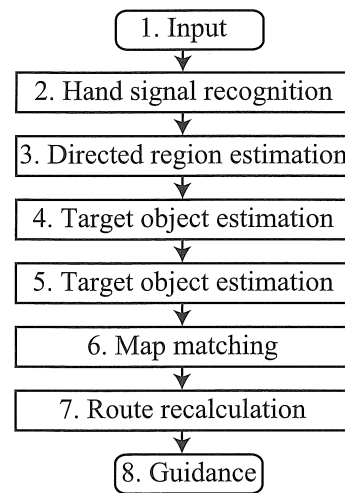


Fig. 2. Processing flow of proposed system

2. 提案システムの概要

本研究では助手席のユーザが操作方法を意識することなく、カーナビの経路案内を補完することを目指す。すなわち人と人とのコミュニケーションで用いるような指差し動作と、従来のカーナビの経路案内を融合させて精度を向上させる。既存のカーナビでは、経路を変更する場合にはタッチパネルやリモコンを操作して新しい経路を探索する必要がある。また、助手席ユーザから経路変更を指示された場合には、事前に示された経路から外れた後で新しい経路が再探索される。一方、提案システムでは、助手席ユーザの指示動作を認識して、逐次新しい経路を探索して提示する。そのため、より直感的でスムーズな経路案内が実現できる。図1に本研究の概念図を示す。本研究の完成により、建物の影に隠れた道路や複雑な交差点においても、助手席のユーザが示す対象物や経路を考慮することにより適切に経路案内をすることが可能となり、例えば目的の交差点を通り過ぎてしまうなどといったミスが低減されることが期待される。また、近年ではメガネ形状のヘッドマウントディスプレイや、ヘッドアップディスプレイとしてのフロントガラスなどに対して、シースルーで経路案内画像を描画する研究が進められており、本研究成果を応用することも可能である。本研究では音声認識、表情認識などを考慮することにより、ユーザとインタラクティブに伝達可能な知的な経路案内システムの構築を目指す。

3. 処理の流れ

提案システムでは、助手席ユーザの指差し動作と風景中の対象物体を認識するために図1に示すように、車内用カメラ(図1中A)と車外用カメラ(図1中B)を用いる。それぞれのカメラで撮影された画像をCV技術により認識し、認識結果を経路案内に反映させる。以下に提案システムの処理の流れを示す(図2)。

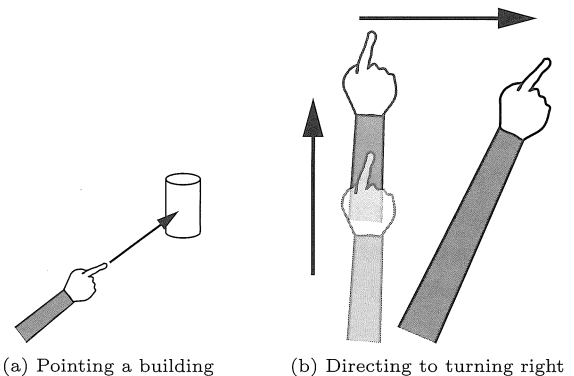


Fig. 3. Examples of hand signals

1. 画像の入力
2. 指差し動作の認識
3. 指示領域の推定
4. 対象物体の推定
5. 地図情報とのマッチング
6. 経路再計算
7. 経路案内をユーザに提示

ここで図中の番号と以下の番号は対応している。これまでに1~3について実験を行ってきた。以下に対象とする指差し動作とその認識方法、そして指示領域の推定について述べる。

3.1 対象とする指差し動作

車内用のカメラで助手席ユーザを撮影し、CV技術により指差し動作を抽出・認識する。経路案内における指差し動作は、図3に示すように、用途によって使い分けられる。図3(a)のような指差し動作の場合、前方の建物や看板、もしくは方向を示し、図3(b)の場合では、経路もしくは対象物体の形状を示すことが多い。上記の指差し動作の種類を場合分けした場合、以下のように分類できる。

1. 対象物体がユーザから目視できる状態で対象物体を示す
2. 対象物体がユーザから目視できる状態で経路を示す

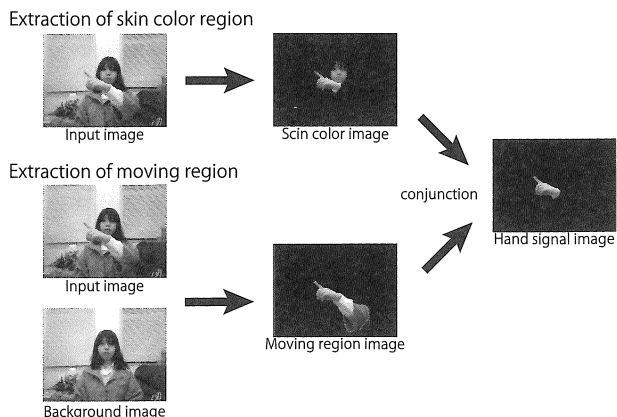


Fig. 4. Extraction with color image

3. 対象物体がユーザから目視できない状態で経路を示す
4. 対象物体がユーザから目視できない状態で目的地の方向を示す

本論文では簡易実験として、視界風景に対象物体が存在することを前提とした指差し動作 (1) を対象とする。

3.2 指差し動作の認識

指差し動作の認識として、我々が提案したカラー画像を利用した手法 [2] と、深度画像を利用した提案手法について述べる。

3.2.1 カラー画像を利用した手法 [2]

指差し動作を認識するために、まず動画中の手の領域の抽出を行う。手領域の抽出には、手の肌色成分と動作部分に着目する。カメラから得られたカラー画像を利用した手領域抽出手法の大まかな流れを図 4 に示す。まずカメラで読み込んだ RGB 画像を HSV 画像に変換し、色相に基づく領域抽出とモルフロジー処理により肌色領域を抽出する。このとき、肌色領域は手だけでなく、顔などの領域も含まれる。一方で、動作領域は、動画中の現在のフレームを X_n としたとき、過去のフレームからの動的背景画像 F_{n-1} との背景差分法により求める。ここで時間的重み α を考慮した動的背景画像 F_n の生成法を次式に示す。

$$F_n = \alpha X_n + (1 - \alpha) F_{n-1} \quad (1)$$

そして、肌色領域と動作領域の積集合を指差し動作中の手

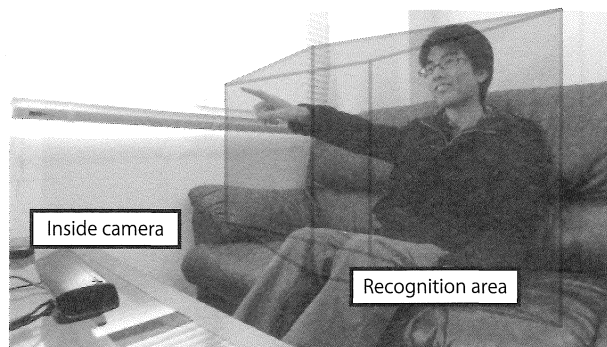


Fig. 6. Limit of recognition area with depth value

領域とする。

$\alpha = 0.08$ としたときの従来手法の結果を図 5 に示す。従来手法では背景画像を動的に更新することによって、車内で移動している助手席ユーザの指差し動作の手領域のみを抽出することができた。一方で車内外の色値が肌色に近い場合、図 5(c) のように十分な抽出精度が得ることができなかった。そこで次節に示すように色情報に依存しない深度情報に着目する。

3.2.2 深度情報の利用した提案手法

肌色や車外風景と手の色値が近い場合による誤検出を防ぐために、カメラとユーザの距離 (深度) 情報を利用する。深度情報の取得には Microsoft 社から発売されている Kinect [8] を用いる。

まず助手席ユーザの指差し動作から「指先座標」を求める。指差し動作はカメラとユーザ身体の間で行われることを仮定し、認識範囲を図 6 のように設定する。認識範囲は、画像中の指差し動作が行われる上下左右領域に範囲を限定する。さらに、深度情報を用いて、認識範囲を車内カメラからの一定距離内に限定する。これは、画像中で指差し動作が移動する範囲と、車内カメラから一定距離の範囲に基づいて決定される四角錐台となる。指差し動作中の指や手の三次元領域は、深度情報を用いて抽出されている認識範囲内の物体とする。そこで、認識範囲で抽出された三次元領域を二次元画像に投影し、「指先座標」を求める。指先座標は、車内カメラから最も近い点が指先となることが多いが、指差しの手の向きによっては指先以外の点が車内カメ

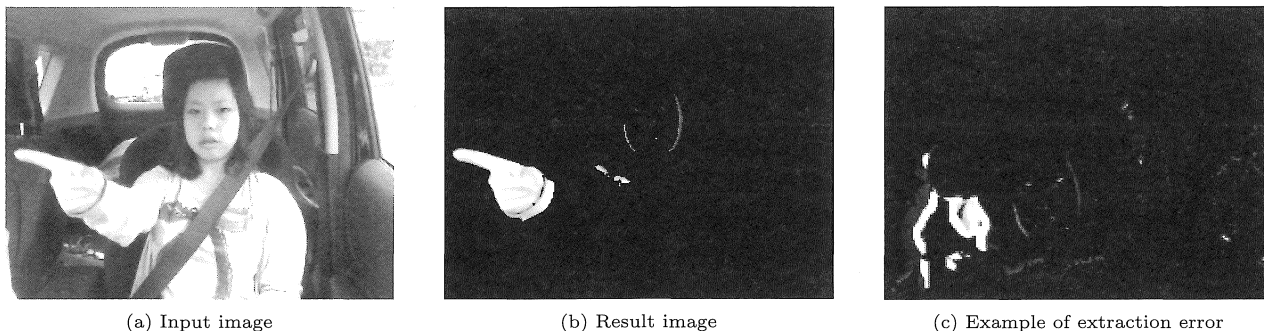


Fig. 5. Extraction result with color video by the former method

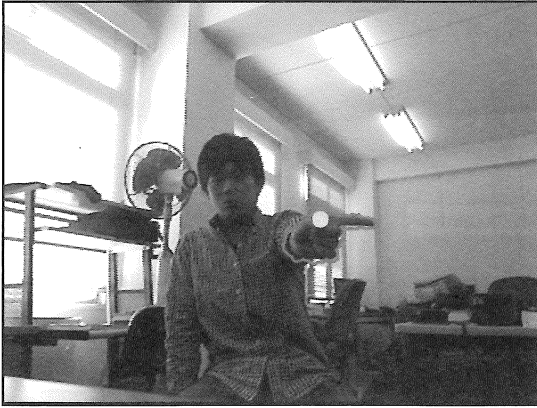


Fig. 7. Example of the nearest point (green) and contour (red)

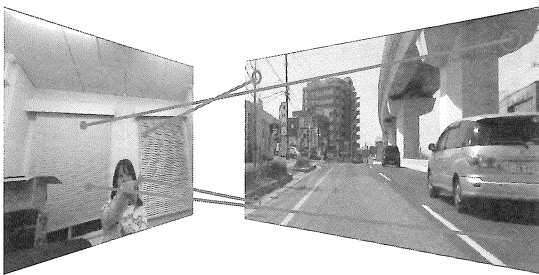


Fig. 8. Camera calibration

ラから最も近くなる場合もある。図7にカメラから最も近い点(緑)と指先が一致しない例を示す。ここで赤色の輪郭は二次元画像に投影された領域の輪郭を示す。そこで本研究では、領域の輪郭形状に着目する。輪郭形状が円に近い場合、指差しはカメラ方向に向いており、円以外の形状の場合、ユーザに対して横方向に向いていることが多いということが事前実験より確認している。輪郭形状が円に近いかを示す円形度 q を円の面積 S 、領域の周囲長 L を用いて次式に示す。

$$q = \frac{4\pi S}{L^2} \quad (2)$$

q が1に近づく程円に近いので、しきい値 t を用いた式(3)で判別する。

$$|q - 1| > t \quad (3)$$

領域の輪郭の形状が円に近ければ車内カメラから最も近い点を、それ以外ならば輪郭の最長距離を持つ2点を求め、そのうち、車内カメラから近い点を「指先座標」とする。

3.3 風景中における指示領域の推定

車外用カメラで前方の風景を撮影し、CV技術により対象物体を抽出する。その前処理として、認識された指先座標により、風景画像中における対象物体が存在する可能性が高い領域を特定して、検索範囲を限定することで対象物体の抽出精度を向上させる。

手順としては指差し動作認識で得られた「指先座標」を

用いて、車外カメラで撮影された風景中における「指示座標」を求める。助手席ユーザに事前に4点の指差し動作を行い、 3×3 の射影変換行列を求め、指先座標を射影変換することで指示座標を得る(図8)。指示座標を中心とする一定範囲内の領域を「指示領域」とする。指示領域を図9に示す。ここで指示領域は指差しの対象となる物体を探索する範囲を限定するために用いる。領域サイズは、風景画像の縦横長さをそれぞれ $\frac{1}{N}$ したものとする。

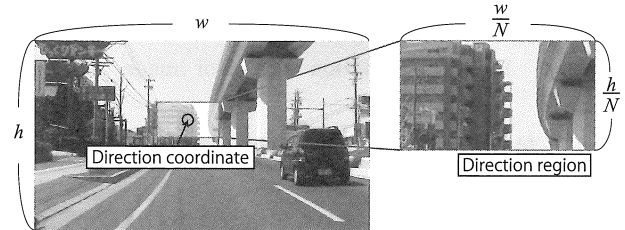


Fig. 9. Size of direction region

4. 実験と考察

4.1 指差し動作認識による指示領域の評価

車外風景をCG (Computer Graphics) 技術で再現し、指示領域の推定についての実験を行った。被験者にはCG画像中の対象物体の指差ししてもらった。計算環境として、Mac OS X, ver. 10.7, 2.3GHz Intel Core i7, 8GBメモリを用いた。画像サイズは車内カメラを640px×480px、風景画像を1600px×900pxとした。深度情報の取得にはKinectを用いる。認識範囲をKinectから0.5~1.1mに限定し、上下左右をそれぞれ入力画像における0px, 80px, 80px, 80pxを除いた範囲とした。円形度の閾値と指示領域の分割数は、先行実験により $t = 0.67$, $N = 5$ とした。

実験の結果を図10, 11に示す。図10(a)の赤円は図10(b)の4点の赤円を指差した際の点である。その4点を基に射影変換することで図10(b)に「指示座標」を緑円に示した。被験者が指差した「指示領域」(図10(c))を結果として判断し、「指示領域」に少しでも対象物体が入っていれば成功、対象物体が入らなければ失敗とした。実験は7人にそれぞれ1回のキャリブレーションを行い、その後2箇所、図10(b)における画像中の中心部、橙と青の看板を指差すように指示した。実験の成功率を表1に示す。橙の看板を指差してもらった場合、71%の成功率であった。しかし、青の看板を指差しをしてもらった場合の成功率は43%と低いことが確認された。その理由として、指差し動作による指示には指先座標と視点の位置が関係しており、キャリブレーション時の補完が線型ではないためであると考えられる。すなわち、横方向中央に存在する青色看板の場合指が遮蔽物になるため、キャリブレーション時の中央部分より少しずらすという傾向がある。今後の課題として視点の位置を考慮

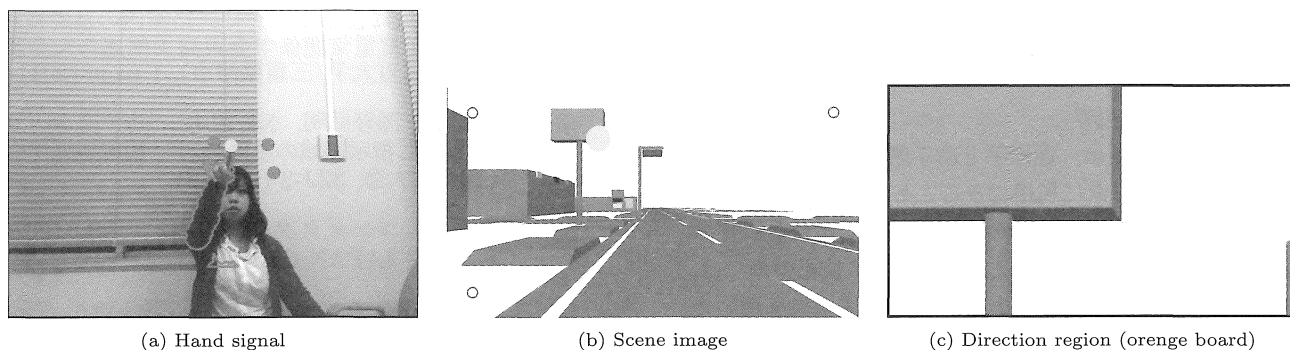


Fig. 10. Experimental Result for orange sign board

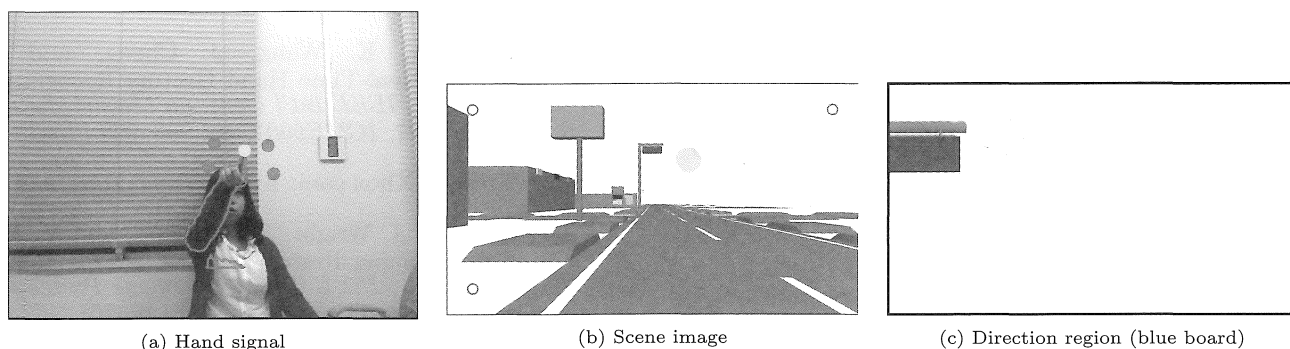


Fig. 11. Experimental Result for blue sign board

Table 1. Evaluation of proposed method

Signboard color	Oreng	Blue	Total
Examinee	7	7	14
Success	5	3	8
Success rate	71	43	53

した手法の構築が必要である。

4.2 課題とその対策

提案システムを現実の車載機に実装するために、様々な課題が予想される。予想される課題とその対策方法を示し、今後の指針とする。

4.2.1 車内用カメラ

車内用カメラは図 12 に示すように、設置位置について考慮する必要がある。今回の実験では、図 12 における A の位置で実験を行ったが、カメラの画角、ユーザの手の位置によっては、図 13 のように指先座標が画像外に存在する可能性がある。特に車載システムではコストの制約を考慮しなければならないため、十分な画角や解像度を持つカメラを使用できる可能性は高くはない。複数台のカメラや超解像度技術 [9] を用いて推定精度の向上を図る。

4.2.2 指差し方向の推定精度の向上

提案した手法では、指先座標のみを利用しており、ユーザの目の位置、指や腕の角度を考慮していない。また、ユーザによっては同じ対象物を指示するように依頼したとしても、同一人物においても指示座標が同一でない場合がある。ユーザの目や指や腕の角度を考慮した推定法の検討につい

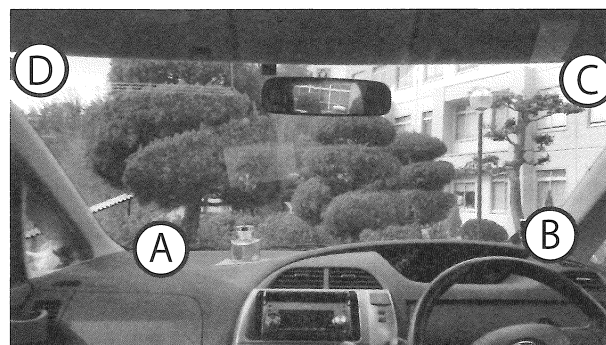


Fig. 12. Camera installation position



Fig. 13. Hand signal outside an image

て今後の課題である。

4.2.3 ユーザによって異なる地図縮尺の考慮

3.1 節で述べたように、助手席ユーザの意図する地図縮尺は一定ではない。すなわちユーザが目視できる対象物の

場合は提案手法を利用できるが, 目視できない場合は, 地図情報を利用した手法の利用について検討する。

4.2.4 音声認識の利用

本システムの利点は, 運転手と助手席のユーザが自然言語で会話しているときに利用できることである。そのため, 会話中のジェスチャーと目的の交差点を示す指差し動作の判定を行う必要がある。本論文で示した CV 技術のみでは, 十分に判定することは難しいため, 音声認識結果を用いた判定は今後の課題である。音声認識が可能になれば, 図 3 に示したような指差し動作自体の判定も可能になることが期待される。

4.2.5 交差点前の車線変更

道路交通法 [6]では, 右左折の場合, 30m よりも手前で車線変更する必要がある。そのため提案システムでは目的の交差点との距離が 30m だけでなく, 経路再探索結果を運転手が知覚するための時間を考慮した手法を検討する必要がある。

4.2.6 指示領域のサイズの動的変更

現在の指示領域のサイズは簡易実験として固定している。しかし移動中の風景において直進方向では遠方になり, 移動方向に対して垂直方向程近い。そこで画像中の位置によって指示領域のサイズを動的に変更することを検討する。

5. おわりに

本論文では, 指差し動作認識と風景認識に基づくカーナビ経路案内について述べた。指差し動作の認識には色値の影響を低減するために車内用カメラからの深度情報を用いて指先座標を抽出した。また, 指先座標を用いてキャリブレーションを行い, 風景中の指示座標を推定した。推定された指示座標から指示領域を求め, 対象物体が存在する可能性が高い領域を推定した。指示領域の推定実験では, 精度が 53% と高くないことが確認された。これは指先座標のみを利用しており, ユーザの目の位置や視線, 腕の角度などを考慮していないことが原因であると考えられる。

また, 提案システムを現実の車載機に実装するための課題をまとめ, 今後の指針を示した。試作システムでは映像のみを対象としているが, 音声認識や地図情報の利用などを考慮して対応する必要がある。今後の課題として, 列挙した課題の解決法の提案及びシステムへの実装である。

文 献

- [1] 澤野弘明, 岡田稔: “車載カメラによる実時間画像処理とその AR 技術に基づく表示方式によるカーナビへの応用”, 芸術科学会論文誌, Vol. 5, No. 2, pp. 57–68 (2006-6)
- [2] 白根愛佳, 澤野弘明, 水野慎士: “指差し動作認識及

び風景認識によるカーナビ経路案内の基礎検討”, 平成 23 年度電気関係学会東海支部連合大会講演論文集 (三重大学/三重県津市), 1 page (2011-9)

- [3] 白根愛佳, 澤野弘明, 水野慎士: “助手席ユーザによる指差し指示領域の基礎検討”, 情報処理学会第 74 回全国大会, 3ZJ-2 (2012-3)
- [4] OKI セミコンダクタ株式会社: “ML512010”, CEATEC JAPAN (2010-5)
- [5] Audi: 2012 International CES <http://www.cesweb.org/>
- [6] 道路交通法: <http://law.e-gov.go.jp/htmldata/S35/S35H0105.html>
- [7] Y. Fang, K. Wang, J. Cheng, and H. Lu: “A Real-Time Hand Gesture Recognition Method”, *IEEE Int'l Conf. on Multimedia and Expo 2007, ICME2007*, pp. 995–998 (2007)
- [8] Kinect - Xbox.com: <http://www.xbox.com>
- [9] J. Yang, J. Wright, T.S. Huang, and Y. Ma: “Image Super-Resolution Via Sparse Representation”, *IEEE Trans. on Image Processing*, Vol. 19, No. 11, pp. 2861–2873 (2010)