

科学研究費助成事業 研究成果報告書

平成 28 年 6 月 6 日現在

機関番号：54501

研究種目：若手研究(B)

研究期間：2013～2015

課題番号：25730066

研究課題名(和文) スマートフォンと光学透過HMDによる高精度で安全なARナビシステムの開発

研究課題名(英文) Development of an accurate and safe AR navigation system using a smartphone and an optical see-through HMD

研究代表者

新井 イスマイル(Arai, Ismail)

明石工業高等専門学校・その他部局等・准教授

研究者番号：60512572

交付決定額(研究期間全体)：(直接経費) 3,200,000円

研究成果の概要(和文)：歩きスマホによる事故を軽減するために、最初のナビゲーションの設定は操作性の良いスマートフォンで行い、その後のナビゲーションを光学透過HMDで行うシステムの各種機能を開発した。光学透過HMD上で視認性高くナビゲーション情報を提示する色彩調整機能と、高精度な位置・姿勢推定を実現するための地磁気フィンガープリンティングの改良およびCNNによる画像マッチング性能の検証を行った。

研究成果の概要(英文)：To realize an accurate and safe pedestrian navigation, we developed parts of AR navigation system which uses a smartphone to setup navigation with rich control interface at first and then uses an optical see-through HMD to safely display navigation information. We conducted three experiments as follows. 1) Developing an automatic color adjustment function to clearly display navigation information on the HMD. 2) Improving geomagnetism fingerprinting for accurate indoor positioning. 3) Verifying CNN (Convolutional Neural Network) for accurate scene matching.

研究分野：情報学

キーワード：ユビキタスコンピューティング

1. 研究開始当初の背景

AR による実世界コンテンツ提示システムは検索エンジンや Web2.0 に続く革新的な情報共有基盤として期待されている。街中に刻まれた実世界コンテンツを見て街の理解を深めたり、災害時の避難誘導ができたり、さらには異なる言語圏で母国語によるナビゲーションを享受できたりすれば、ユーザに広く受け入れられる持続性の高いサービスになる。AR の実現にあたっては、現実世界の映像にズレなく実世界コンテンツを重畳表示するための「高精度な屋内測位・姿勢推定」および、ナビゲーション等のリアルタイムサービスを楽しむ中、歩きながらも自然に、かつ他の通行人と衝突しないよう、安全にシステムを利用するための「AR を安心・安全に利用し続けられるユーザインタフェース」の技術確立が鍵となる。

前者の測位・姿勢推定は、屋外においては建物等実世界コンテンツ投影対象の密度が低く、GPS や地磁気センサも補正技術との組み合わせで良好に動作するため問題は少ないが、屋内では実世界コンテンツを重畳表示したい「モノ」の密度が高く、また GPS 等が利用できないことから、別個の測位・姿勢推定技術が求められる。理想とする数 cm～数 10cm 程度の測位・姿勢推定誤差を達成する手法として、2次元マーカを環境側に貼付する手法や、超音波・高速点滅 LED 等のビーコンを敷設するものがあるが、インフラ構築コストが高い。一方でインフラ構築コストの低い既設無線 LAN 基地局の観測情報を活用したものは精度が数 10m と低く AR に適さない。

後者の「AR を安心・安全に利用し続けられるユーザインタフェース」については、現在普及しているスマートフォンでは対応が難しい。スマートフォンを見ながら歩くことで周辺への注意が薄れ、混雑地域での衝突事故や電車のホームでの落下事故が問題となっている。スマートフォンの画面が小さいことで凝視が必然となること、またその視線方向がその人の進行方向と異なることに問題がある。

2. 研究の目的

地下街等の広域な屋内公共空間を歩行中のユーザに、直感的で安全な AR によるナビゲーションを提供するために、スマートフォンと光学透過 HMD (Head Mount Display) を連携したシステムを開発する。適切な位置にセンサを配置することで「高精度な屋内測位・姿勢推定」を実現し、適切な位置にナビ情報を表示することで「AR を安心・安全に利用し続けられるユーザインタフェース」を実現する。

3. 研究の方法

図 1 の計画に従って、研究開発を進めた。

各項目の詳細は下記の通りである。

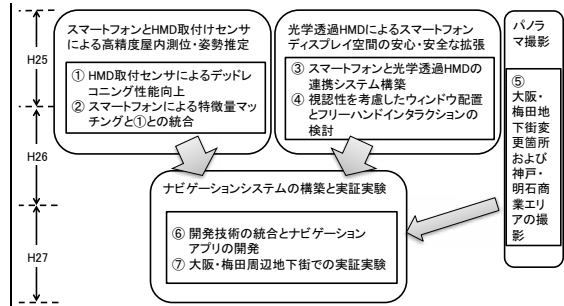


図 1. 研究計画

- (1) HMD 取付けセンサによるデッドレコニング性能向上
HMD に取付けるセンサに役割を持たせることで、デッドレコニング性能を向上させる。センサ値の処理は全てスマートフォンで行うことを想定しているが、HMD 側センサの生データを出力するのが困難な場合は、HMD 側にマイコンを取り付けて加工後のデータをスマートフォンに送信するシステムを構築する。
- (2) スマートフォンによる特徴量マッチングと①の統合
過去に開発した屋内測位・姿勢推定システム (Wi-Fi フィンガープリンティングと地磁気・加速度による方位推定、画像特徴量マッチングを組み合わせた高精度屋内測位システム) と①のデッドレコニング技術を組み合わせて、高精度に連続した屋内測位・姿勢推定を実現する。屋内測位手法として地磁気のフィンガープリンティング手法が注目を集めており本研究のフィールドにも適用可能性があるため基礎研究を行いつつ、画像特徴量マッチングの精度改善を CNN (Convolutional Neural Network) の適用によって可能かどうか検証する。
- (3) スマートフォンと光学透過 HMD の連携システム構築
光学透過 HMD の画面をスマートフォンから操作できるインタフェースを検討する。
- (4) 視認性を考慮したウィンドウ配置およびフリーハンドインタラクションの検討
光学透過 HMD に表示する映像が歩行中に邪魔にならないようなナビゲーション情報提示手法を検討する。またフリーハンドインタラクションの手法として騒音中の音声認識の精度向上を図る。
- (5) 大阪・梅田周辺地下街変更箇所および神戸・明石商業エリアの撮影
所有している大阪・梅田周辺地下街の店舗入れ替え等に対応する。可能であれば他の商業施設の実験も行う。
- (6) 開発技術の統合とナビゲーションアプリケーションの開発
システムを統合して実フィールドでの実用可能性を検証する。開発状況によって

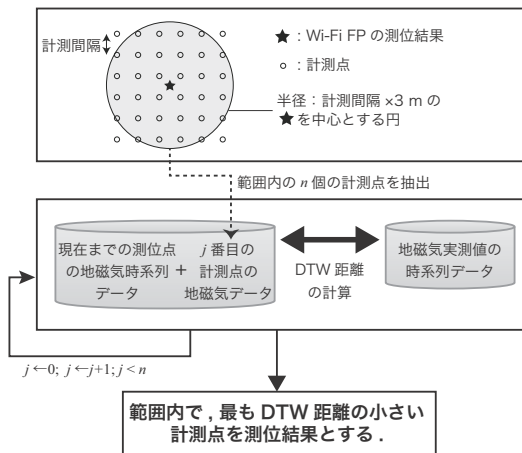
は基本的な性能評価のみ行う。

4. 研究成果

計画(1), (2)については相互補完型 Wi-Fi・地磁気フィンガープリンティングと、CNN を応用した画像シーンマッチングについて成果を得た。CNN 応用に際して(5)の再撮影を大阪・梅田周辺地下街で行った。また、(3)については背景色に基づくビューマネジメントを、(4)については騒音下における音声認識の精度改善について検証を行った。(6)については上記要素を統合して実証実験するには至らなかったため、上記実験の性能評価をもって完了とした。

4. 1. 相互補完型 Wi-Fi・地磁気フィンガープリンティング

設備負担のない屋内測位手法で、スマートフォン搭載センサを駆使するものとして Wi-Fi・地磁気フィンガープリンティング (FP) の双方を組み合わせた手法 WaMM が提案(図 2) されている。既存研究では Wi-Fi FP の測位結果周辺で地磁気 FP を行うことで測位する。しかし提案されている手法では一方の FP が良い精度であっても、もう一方の FP の精度に測位結果が大きく影響されるという性質がある。



この問題点を解決するために、観測した Wi-Fi アクセスポイント (AP) の BSSID 観測状況に応じて制限したエリア内で地磁気 FP を行った結果と、Wi-Fi FP の結果を、Wi-Fi FP の信頼度で加重平均することで、安定した精度を得る手法 (図 3) を提案した。

評価の結果、測位誤差の平均値 6.95m、中央値 3.48m、測位失敗率 0%となった。また実験環境上の AP の BSSID を無作為に削減した時の、測位誤差が 5m 以内に収まる確率について、Li らの手法は 28%に対し、提案手法は 48%を達成し、Wi-Fi FP の高い精度を維持したまま、Wi-Fi FP で測位できないエリアも安定した精度で測位できた。(図 4)

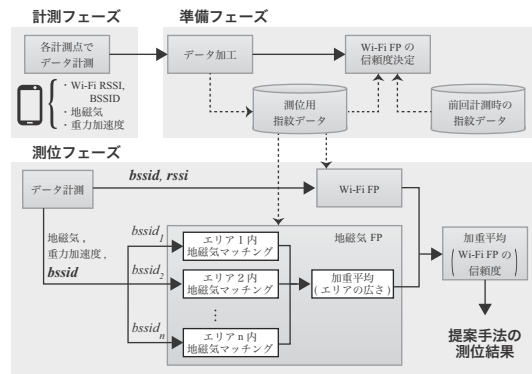


図 3. 提案する相互補完型 Wi-Fi・地磁気 FP

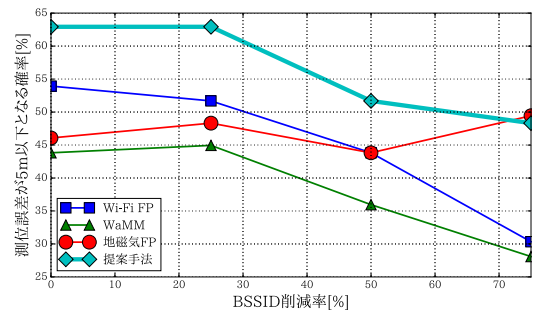


図 4. 測位誤差が 5m 以下となる確率

4. 2. CNN による画像シーンマッチング

画像とその他のセンサを組み合わせた位置・姿勢推定手法が種々提案されているが、画像による位置推定の性能や画像の特徴点の量による処理速度の性能が課題となっている。そこで、CNN (Convolutional Neural Network) を用いることで、高精度な画像分類による位置推定の精度向上や、特徴点の量に影響されないことによる処理速度向上が期待できるシステムを提案する。CNN は近年、画像認識・分類のコンテストで良く用いられているが、実環境では使用例が無いため、実験場所となる大阪・梅田周辺地下街の看板画像の分類結果から、実際にシステムに利用できるかどうかの評価を行った。

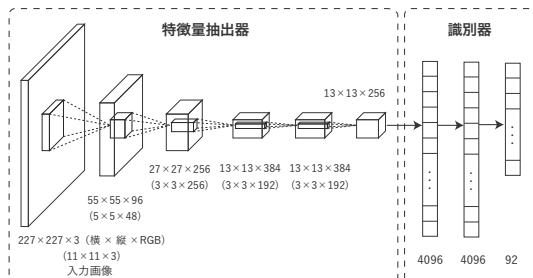


図 5. CNN の構成図

実験で用いた CNN の構成図は図 5 の通りになっており、既存の画像分類で用いられていたパラメータ設定を踏襲している。大阪・梅田周辺地下街で撮影した看板画像が 92 種類あるため、最終的な分類結果が 92 となるように構成されている。



図 6. 訓練画像

訓練画像は一般的に 1 分類あたり約 1,000 枚必要とされているが、人通りの多い地下街にて上下左右に満足な種類の角度でじっくり撮影することが困難なため、フル HD のハンディカムで 1,000 フレーム程度を撮影できる歩行速度で撮影しながら地下街を巡回撮影した映像を利用した。切り出した看板画像は図 6 のようなもので鮮明である。

実験結果より、訓練画像と同じ条件のテスト画像に関しては約 93 [%] という正答率を得ることができた。別途用意したメガネ型端末 (30 万画素) のクエリ画像は 10 [%]、スマートフォン (800 万画素) のクエリ画像は 59 [%] という結果となった。それぞれ、成功例と失敗例を並べたものを図 7、図 8 に示す。訓練画像の撮影角度の種類の貧困さにより訓練画像のクロス評価と実利用に近い状況の性能に差が開いた結果となった。また、クエリ画像の画質も性能を大きく左右するため現場の HMD 搭載カメラでは実用に至らないことが分かった。ただし、HMD に搭載に当たる制約を強く受けているわけではないため、需要に応じて HMD 搭載カメラの性能を上げることは難しくないため、今後の実用性は向上すると考えられる。



図 7. HMD 画像の認識例



図 8. スマートフォン画像の認識例

4. 3. 背景色に基づくビューマネジメント

Google Glass を皮切りに HMD の普及が期待された際に様々なアプリケーションが考案されたが、既存のインターフェースでは表示情報が多すぎて背景が見えない、背景に埋もれてしまっていて情報が見えない、などのビューマネジメントの問題が生じる。本研究で使用する光学透過 HMD は特に上記の問題が目立つ。

そこで、背景色によって表示情報の色を動的に変化させる事によるビューマネジメントを設計、実装した。提案するビューマネジメントの概念図を図 9 に示す。表示色を背景画像に合わせて決定してコントラストを確保する。アルゴリズムは図 10 の通りである。基本的には RGB 値を反転することで高コントラストを得るが、灰色等の無色中間の色の場合に輝度の調整を追加している。



図 9. 提案するビューマネジメントの概要

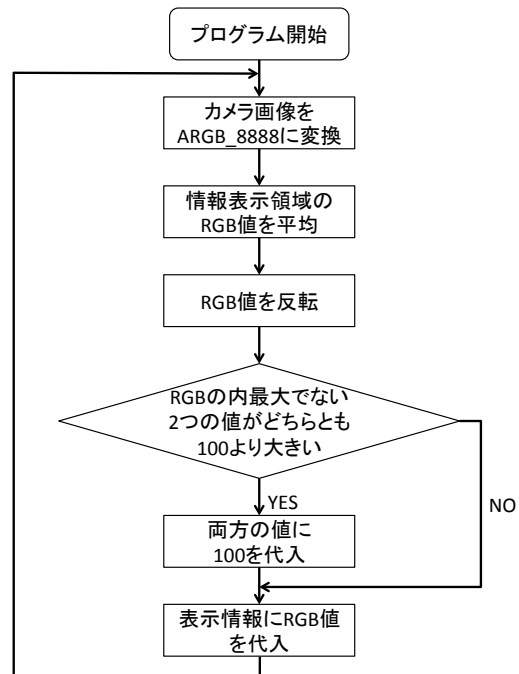


図 10. 提案アルゴリズム

光学透過型 HMD を用いて評価実験を行ったところ、背景色と表示情報の色が近くなってしまう事による視認性の低下を防ぐ事ができるといった結果が 5 人の被験者のアンケートによって得られた。

4. 4. 騒音下における音声認識の精度改善

HMD によるナビゲーションを受けている際に何かしら操作をしたい場合に、本研究の想定では立ち止まってスマートフォンをする事になるが、単純な操作の場合は音声認識を用いる手法も考えられる。しかしながら、立ち止まって操作ができないような人通りの激しい環境は騒音が音声認識のノイズ源となる。このような環境では骨伝導マイクを用いて発話者以外の音声を取り込まないようにする手法が適していると考えられるが、骨伝導音は耐ノイズ性の代わりに周波数帯域が狭いため、新たな音響モデルとその認識器を構築する必要がある。本研究では通行人の声やアナウンスの声を除去するタスクは難易度が高かったため、第一段階の別のアプリケーションとしてバイク走行中の騒音下の音声認識性能の改善に取り組んだ。

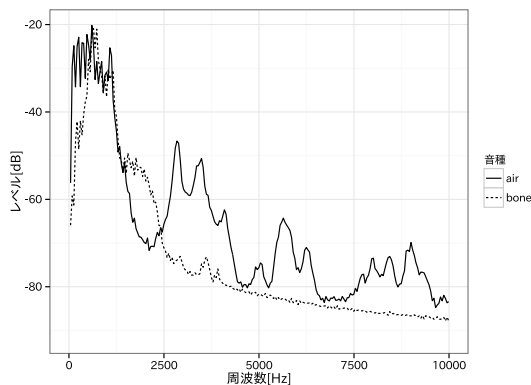


図 11. 音声「あ」のスペクトル

「あ」という音に対する骨伝導音と気道音のスペクトルの違いを図 11 に示す。2.5kHz 以降は大きく減衰しており、これは「あ」という音声を認識する上で重要な第 3 フォルマント (3.0kHz 辺り) を失っているため、従来の気道音による音声認識を適用することが困難となる。既存研究に雑音を含んだ気道音のモデル構築と音声認識の研究と、骨伝導マイクの第 3 フォルマント周辺を増幅して気道音モデルに適用した研究があるが、前者は全ての雑音に対応できない問題、後者は未だに再現が不十分である問題が残っている。そこで、雑音を含んだ (とはいえ骨導マイクのため大幅にノイズ削減されている) 骨伝導音を直接学習した音声認識システムを構築した。

兵庫県明石市周辺にて走行中に目的地を変更するシナリオで音声認識を行った。音素のマッチングにおける単語の認識率は図 12 の結果となった (従来手法はノイズ込みの気道音の学習、提案手法 1 はノイズなし環境で収録した骨伝導音の学習、提案手法 2 はノイズ込みの骨伝導音の学習)。

実際の音声認識では n-gram によって多少の音素の認識ミスを吸収するため、それを適用した結果、98.75%の認識率を得た。従来手法では 93.75%の認識率だったため 5%の性能向上を達成した。

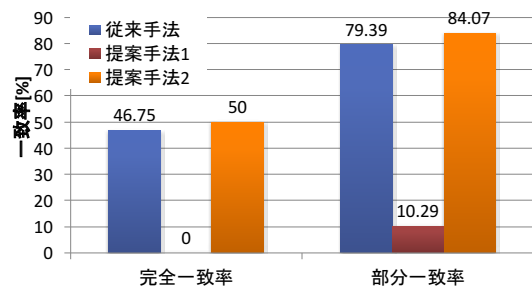


図 12. 音声認識評価の結果

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

[学会発表] (計 5 件)

- (1) 東 和樹、新井 イスマイル、“相互補完型 Wi-Fi・地磁気フィンガープリンティング手法の開発”、電気学会関西支部平成 26 年度高専卒業研究発表会、2016 年 3 月 5 日、中央電気倶楽部。
- (2) 東 和樹、新井 イスマイル、“相互補完型 Wi-Fi・地磁気フィンガープリンティング手法の開発”、情報処理学会 DPS 研究会、2016 年 3 月 3、4 日、明治大学。
- (3) 立川 剛至、新井 イスマイル、“センサ・カメラ併用型屋内測位・姿勢推定システムへの CNN 応用の提案”、インターネットコンファレンス (IC2015)、2015 年 10 月 13、14 日、(公) 兵庫産業活性化センター ビジネスプラザひょうごホール。
- (4) 東 和樹、新井 イスマイル、“相互補完型 Wi-Fi 地磁気フィンガープリンティング手法の提案”、インターネットコンファレンス (IC2015)、2015 年 10 月 13、14 日、(公) 兵庫産業活性化センター ビジネスプラザひょうごホール。
- (5) 富士原 匡隆、新井 イスマイル、“光学透過型 HMD を用いた AR ナビにおける背景色に基づくビューマネジメント”、情報処理学会 UBI 研究会、2014 年 3 月 14、15 日、慶應義塾大学日吉キャンパス。

6. 研究組織

(1) 研究代表者

新井イスマイル (ARAI ISMAIL)

明石工業高等専門学校・電気情報工学科・准教授

研究者番号：60512572