

障害者自立支援機器等研究開発プロジェクト

画像・GPS等のセンサ統合による日常利用可能な
屋内外視覚障害者歩行支援システムの開発に関する研究

平成21年度 総括・分担研究報告書

研究代表者 静岡県立大学 石川 准

平成 22 (2010) 年 4 月

I. 障害者自立支援機器等研究開発プロジェクト

総括研究報告書

画像・GPS等のセンサ統合による日常利用可能な屋内外視覚障害者歩行

支援システムの開発に関する研究

研究代表者 静岡県立大学 石川 准

研究要旨

本研究の目的は、視覚障害者の移動の自由と移動の安全を各種センサ技術により支援することである。歩行の自由を支援するためには、位置方位計測の精度向上、及び適用範囲の拡大が必要であり、歩行の安全を支援するためには障害物検知技術の向上が求められる。そのために、画像認識、歩行者デッドレコニング、LRFといった要素技術を視覚障害者歩行支援に適用するための応用技術開発を行う。さらに、画像認識、歩行者デッドレコニング、GPS、プレースエンジン、マップマッチング等を統合する「センサ統合技術」を開発する。また音声点字携帯端末上での歩行ナビゲーション情報と障害物検知情報の効果的提示方法を開発する。統合システム実証実験、統合システムモニタ実験を実施し定量的および定性的な性能評価、操作性評価を行う。

研究分担者

青木恭太・宇都宮大学・准教授
亀田能成・筑波大学・准教授
蔵田武志・産業技術総合研究所・主任研究員
興梠正克・産業技術総合研究所・研究員
石川智也・産業技術総合研究所・特別研究員
湯瀬裕昭・静岡県立大学・准教授

A. 研究開発目的

本研究の目的は、視覚障害者の移動の自由と移動の安全を各種センサ技術により支援することである。歩行の自由を支援するためには、位置方位計測の精度向上、及び適用範囲の拡大が必要であり、歩行の安全を支援するためには障害物検知技術の向上が求められる。

そこで、GPSでの正確な測位が困難な高マルチパス環境の都市市街地や、GPS測位が不可能な屋内環境を対象として、画像認識に基づく位置方位推定および歩行者デッドレコニング(PDR)とGPSおよびWi-Fi測位の統合による高精度歩行ナビゲーションの実現をめざす。さらに画像認識とレーザレンジファインダによる衝突、転倒、落下等の事故を防ぐための障害物検知技術の開発も行う。また将来的な共有画像データベース(DB)整備を想定した画像登録方法の

開発もめざす。

そのために以下の個別目標を立てた。

- (1) 画像中の看板、横断歩道等を対象とするランドマークベース画像認識と、画像全体を対象とするシーンベース画像認識技術を開発して位置推定や施設検索を行う。
- (2) 歩行者デッドレコニング(PDR)については、小型PDRセンサモジュールを開発し視覚障害者の歩行パターンに合うように改良する。
- (3) 上記の各技術を用いて、GPSやWi-Fi測位を補正するためのセンサ統合技術を開発する。
- (4) 画像中の看板、横断歩道等を対象とするランドマークベース画像認識と、画像全体を対象とするシーンベース画像認識を組み合わせた位置方位推定技術を実現するためには、画像データベースが必要となる。共創の概念に基づき、その整備コストを視覚障害者だけではなく健常者等と分散できるような枠組みを導入するために、画像登録支援ツールを開発する。また、本プロジェクトで開発する視覚障害者歩行支援システムの測位誤差評価を行う
- (5) 画像処理、レーザレンジファインダ(LRF)、PDRを用いた障害物検知技術を開発する。
- (6) カメラ付き音声点字携帯端末のユーザインタフェースを開発し、音声メッセージ、効果

音、点字メッセージ、触覚グラフィックス等を用いた情報提示方法の開発を行う。

- (7) 個別要素技術の実証実験と東京駅周辺の実証コースでの統合モニタ実験を行い、研究結果の性能評価、操作性評価を実証的、定量的に示す。

B. 研究開発方法

まず要素技術開発を行う。画像認識による位置・方位補正のために、ランドマークベース画像認識モジュールを開発する。屋外定型看板の実画像を昼間および夜間に収集して撮像条件および前処理条件を決定する。それとともにシーンベース画像認識モジュールを開発する。事前に経路に沿って登録しておいた画像データ群の中から、現在の画像に最も適合する画像を探索し、その画像の撮影位置情報をもって現時刻の位置・方位推定を実現する。

また PDR による相対位置・方位計測アルゴリズムの視覚障害者の歩行パターンへの対応を行う。加速度、角速度、磁気、気圧の各センサデータを用いて相対移動距離や体の向きを推定する PDR は、従来健常者向けに開発されているため、視覚障害者の歩行パターンでの検証や改良が必要となる。

これらに基づき、GPS、Wi-Fi 測位、画像認識、PDR のセンサ統合や、それらにマップマッチング情報を組み合わせた最適統合技術を開発する。

画像処理による障害物検知、LRF と PDR による障害物検知のための要素技術を開発する。特に前方下向きの障害物検知については、その精度向上のために、PDR から得られる相対運動に基づく LRF 時系列データのスタビライズ機能を開発する。

さらに、音声メッセージ、効果音、点字メッセージ、触覚グラフィックスや点字ブリンク等を用いたプル型及びプッシュ型の効果的な情報提示方法を、センサ特性を考慮しながら開発する。

また、共創的ナビ情報インフラ整備のために、画像 DB サーバを開発し、画像登録支援ツールを開発する。

次に、これらの機能を統合して音声点字携帯端末上にユーザインタフェースを構築する。

カメラ付き音声点字携帯端末、携帯サーバ、GPS 受信機、PDR、LRF モジュールからなる試作システムを準備し、東京駅八重洲地下街および八重洲屋外での統合実証実験およびモニタ実験を行う。

モニタ実験終了後に被験者によるグループディスカッションを行い、当事者の評価や要望を聴取する。

(倫理面への配慮)

システム評価のために東京駅八重洲地下街、中央通り、さくら通りにて統合モニタ実験（2 回、のべ 10 人）を行った。この統合モニタ実験を行うために、静岡県立大学内の研究倫理審査委員会に『センサ統合による日常利用可能な屋内外視覚障害者歩行支援システムの実証実験』についての研究倫理審査申請を行い、承認を受けた。

モニタ実験においては研究倫理審査申請書中のモニタ評価の倫理面等に関するすべての事項を厳格に遵守した。

C. 研究開発結果

研究により得られた成果は以下である。

(1)-I. ランドマーク認識による位置推定: ランドマーク DB に登録されたランドマークについて、接近した場合に認識可能となった。ランドマークを認識した場合に、方向および距離の推定が可能となった。

(1)-II. シーン認識による位置推定: 東京駅周辺の実証コース（1km ほど）を想定し、経路上の 200～500 地点程度で過去に撮影された画像群を用いて、現在の入力画像（3 枚使用）から最も尤度が高い画像を参照しその位置を返す機能を実現した。問い合わせから 3～5 秒程度で検索結果を回答する。また、位置推定精度に対する指標化も可能であり、それを信頼度の形で合わせて回答できる。

(2)PDR: 簡易防水小型 PDR センサモジュールを試作し、実証コースにおいて、おおむね良好な結果を得た。従来は腰部に装着していたセンサを他の装置と共に胸部に装着した場合の検証、さらに、視覚障害者の歩行パターンが PDR アルゴリズムに適用可能かを検証し、歩行動作推定パラメータの変更で対応できる

ことがわかった。

- (3) センサ統合による位置・方位推定精度向上: 画像認識、PDR、GPS、プレースエンジンから位置情報・方位情報を受け取り、一定以上の信頼度の場合に PDR の位置情報・方位情報を補正するという方式で、GPS とマップマッチングより高い推定精度を実現した。
- (4) 画像 DB: 画像 DB に関しては、画像 DB サーバを運用しに画像認識に必要な属性を付加した 5500 枚の画像を収集した。画像登録支援ツールでは、シーン・ランドマーク画像双方の登録、また、登録ユーザ層拡大のために、画像属性の簡易登録と詳細登録を行えるようにした。抽出された画像特徴の可視化も行った。また、各研究分担者やその関係者が試用してそのフィードバックに基づいてユーザビリティを高めた。視覚障害者歩行支援システムの測位誤差評価に関しては、各センサの屋内外での特性の比較を地図上の軌跡、さらには統計的に明確にすることができた。また、センサ統合により各センサ単体よりも優れた測位結果が得られ、屋内外合計約 1 km の実証ルートにおいて、平均 6.7m (標準偏差 5.5) の誤差での測位ができることがわかった。
- (5)-I. 画像による障害物検知: 歩行中に短時間に連続して撮影した画像対を用いて、障害物検知を実現した。検知障害物の方位および距離を推定することが可能となった。
- (5)-II. LRF による障害物検知: PDR による重力方向追跡等に基づく LRF 情報の安定化機能を開発した。
- (6) カメラ付き音声点字携帯端末のユーザインタフェース: ルートガイダンスと障害物検出に関して効果的な情報提示方法を考案した。
- (7) 個別要素技術の実証実験と実証コースでの統合モニタ実験: センサ統合による位置・方位推定の精度向上を定量的に実証した。ガイダンス情報、障害物情報の提示方法についてモニタの多様な意見を収集した。

D. 考察

本プロジェクトは先に示したような成果を上げたが、以下のような今後解決すべき問題、課題も明らかになった。問題の明確化は目標の実現にとって重要であり、これもまた本プロジェ

クトの成果の一つともいえる。

- (1)-I. ランドマーク認識による位置推定: ランドマーク認識系の残された課題は、動作の高速化と高精度化である。複数のプロセスが 1 つのハードウェア上で動作する今回の方式では、系の複雑さのために動作速度の向上が困難である。自律分散系の構成により安価に高性能のランドマーク認識を利用した歩行支援システムを構成することが今後の課題である。また、ランドマークデータの WEB 上情報からの自動収集及び、ボランティアによる寄付映像などに基づく自動ランドマークデータ収集システムの構成が今後の課題である。
- (1)-II. シーン認識による位置推定: カメラ撮像機構の小型化と屋内外等の光量変化への対応 (オートアイリスレンズの導入ないし屋内専用・屋外専用の 2 眼化)、メモリ使用量の削減、参照画像集合の拡充 (同一地点で環境を変えて複数撮影) が課題である。
- (2) PDR: ハードウェアについては、さらなる小型軽量化、低コスト化が求められる。また、手法については、PDR センサモジュール内蔵の各センサのキャリブレーションやユーザごとの歩行パラメータのキャリブレーションの簡易化等が課題である。
- (3) センサ統合による位置・方位推定精度向上: GPS の測位データと PDR 等の他のセンサからの情報を照合して GPS 高マルチパスを検出する技術の開発が求められる。
- (4) 画像 DB: 画像の登録方法としては、登録作業負荷のさらなる低減が求められる。運用上の課題としては、サーバ維持コスト、登録協力者が増加した際の登録情報の信頼性の担保、画像を収集する上でのプライバシー保護等をどのように軽減、解決していくかが課題である。
- (5)-I. 画像による障害物検知: 加速度センサ統合による高精度化と自律分散構成による高速化が今後の課題である。
- (5)-II. LRF による障害物検知: 各センサの非同期処理を考慮して PDR による重力方向追跡等に基づく LRF 情報の安定化性能を向上させ、LRF 情報履歴を用いた地形判定能力を高めることが課題の 1 つである。また、LRF を小型軽量のものにしたため、レーザ光の出力

が十分ではなく、日射のある屋外の黒いアスファルト等の路面では S/N 比が極端に悪くなるため、路面が欠落しているのか反射していないのかの判別がつかない。ハードウェア開発以外でのこの課題の解決方法としては、例えば、レーザ光の照射範囲を制限する代わりに 1 箇所辺りの S/N 比を上げるようなレーザ制御法の導入が考えられる。

- (6)・カメラ付き音声点字携帯端末のユーザインタフェース:視覚障害者の状況は様々であり、すべての視覚障害者にとっての最適な単一のインタフェースはない。多様なインタフェースを実現するための統一的枠組みの開発が今後の課題である。

E. 結論

本プロジェクトは、センサ統合にはマルチパス検出技術が不可欠なことや、画像認識には、処理速度や天候条件依存という解決すべき問題が明らかになったが、センサ統合が高精度の歩行支援を提供できる有望な技術であることを実証した。

モニタ実験のグループディスカッションから視覚障害者の障害物検知への高い期待を知ることができた。障害物検知技術は画像、LRF ともになお技術的課題が残されているが、とくに LRF は、特性上の制約はあるが、処理の速さと情報の精度から、近距離の障害物検知技術として有望といえる。

本プロジェクトは音声点字携帯情報端末へのナビゲーション情報と障害物検知情報を有効に提示する方法を考案できた。とはいえ利用者の特性は多様であり、情報提示は多様な形式に対応する必要がある。

モニタ実験からは多くの知見を得た。とくにグループディスカッションからは多くのことを学んだ。視覚障害者はある決められた場所を安全に歩ければそれでよいと考えているわけではない。自由に歩きたい、楽しく歩きたいと望んでいる。歩行支援技術はそのような希望に応える技術でなければならない。

F. 研究発表

1. 学会発表

- ・ M. Kourog, T. Ishikawa, Y., J. Ishikawa, K.

Aoki, and T. Kurata, Pedestrian Dead Reckoning and its applications, In Proc. ISMAR2009 Workshop: Let's Go Out: Research in Outdoor Mixed and Augmented Reality, http://www.icg.tugraz.at/News/outdoor_ar (2009).

- ・ 査読付き国際会議投稿中 1 件 (採否通知 4 月 7 日予定)

- ・ 電子情報通信学会パターン認識・メディア理解研究会 (2010 年 5 月発表予定)

Ⅱ. 障害者自立支援機器等研究開発プロジェクト

分担研究報告書

1. シーン画像認識に関する研究

研究分担者 筑波大学 亀田能成

研究要旨

カメラで撮影した画像を手がかりに、現在位置を推定する研究を行った。位置推定は、過去に既知の地点から撮影した画像群と比較し、最も類似性が高い画像を検索することで実現する。これは一般画像検索技術を歩行者ナビゲーションに応用したものであり、世界的にもまだ例が少ない研究分野である。歩行案内対象の経路に沿って934枚程度の画像を用意することで、入力画像に対して2秒内で歩行者の位置を求めることに成功した。このとき、撮影状況の異なる画像集合を用意しておくことで、日照状況の変動に対しても対応した。

A. 研究開発目的

視覚障害者の歩行支援において、位置推定は重要な情報である。歩行者の位置推定には様々な手法が考えられる。その中でも、歩行者から見える風景(シーン)を撮影し、その画像を手がかりに現在地を求める方法を研究しシステム開発を行うことが本研究計画の目的である。

視覚障害者の外出の一形態として、目的地や歩行する範囲をある程度計画してから外出に臨むことがある。目的地がわかっていたとしても、その途中の経路に関する知識が不確かであったりする場合、歩行支援システムによってナビゲーション情報を正確に提示することは、視覚障害者への有用な支援となる。

このことは、カメラに依る位置推定方法の開発にあたっては、全く未知の経路で位置推定を行うのではなく、ある程度歩行範囲や経路を予め絞り込んでおけることを意味する。

その一方で、一般に目的地や経路選択は多岐に渡ると考えられるので、カメラ位置推定において準備に手間がかかることは避けるべきである。これは、特に多くの人を通るわけではない主要ルートなどが歩行範囲として想定される場合は重要な要因である。

そこで、本研究では、歩行者の胸部前面に搭載したカメラ(歩行者カメラ)で撮影した風景画像を手がかりに、予め規定された歩行範囲内において現在位置を推定する手法を提案し、シ

ステムの開発を行う。

システムの開発に当たっては、現に歩行中の視覚障害者に支援を行う必要があるため、携帯可能なノートパソコン程度の性能でもリアルタイムに位置推定できることが求められる。本研究計画では、問い合わせ実行時から位置推定結果の提示までにかかる時間を数秒内に抑えることを目標に掲げ、1,000地点程度の経路上の登録位置の中からほぼ2秒内に現在位置に最も近いものを示すことに成功した。

B. 研究開発方法

外出計画に含まれる歩行経路には様々な経路が考えられるため、事前の準備にかかる作業量は小さいほどよい。そこで、本研究では、カメラで撮影する風景画像全体を手がかりに使うことにする。これによって、予め蓄積しておくべき画像については、経路に沿って単に撮影して回るだけで収集することができるようになるので、画像収集に対する作業量は小さくなる。

こうして集められた画像には、撮影位置と撮影方向の情報を予め付与した上で画像データベースに登録しておく。

外出前には、これから歩行する可能性のある経路上の画像を集めて画像集合を生成する。歩行中は、この画像集合に対して、コンピュータビジョンの一般画像認識において用いられる手法を応用して歩行者カメラで撮影された風景画

像をキーとして検索を行い、最も見かけが近い画像候補を得る。この画像候補が入力風景画像と十分に似ていると判断された場合、その画像候補に付与されていた撮影位置と方向を、現在の歩行者の撮影位置と方向として出力する。

以下、手法の実現に当たって、さらに細部検討を行う。

・カメラの取り付け方向

カメラの取り付けに関しては、視覚障害者歩行支援システムの一部として歩行者カメラが取り付けられるということもあって、一定の制約が課せられた。具体的には、歩行者カメラは胸部前面に配置し、歩行方向にほぼ水平に向けられた。注意すべき点としては、歩行者カメラは被験者の胴体にしっかりと取り付けるわけにはいかないため、歩行の揺動に合わせてカメラの光軸はかなり揺れることが挙げられる。そのため、後述する画像照合手法では、これらの揺れに対処できることが必須である。

・カメラ撮像機構

それほど高速な移動を伴うわけではないものの、歩行に基づく揺動の中で撮影を行うため、グローバルシャッターは正確な像を得るために必要である。シャッタースピードについても情景が明るい限り、高速にしたほうがぶれなく正確な像を得ることができる。解像度に関しては、予備実験を通して、シーン画像ではそれほど解像度が不要であるとの結論が得られていたが、歩行者カメラを用いる他のモジュールとの兼ね合いもあり、高解像度カメラを用意することになった。これらのことから、Imaging Source社のDFK31AU03（水平1024画素、垂直768画素、最高で秒30枚撮影可能）を選定した。

・レンズ

歩行に基づく揺動があることから、カメラの画角は広いほうが好ましい。ただし、歩行者カメラを用いる他のシステムとの関係上、水平画角61.6°、垂直画角48.3°になるFujinon TF4DA-8を最終的に選定した。

なお、屋内から屋外にまで撮影が渡るため、レンズ自体にオートアイリス機能があることが

望まれたが、ハードウェア選定の制約から、本研究計画ではその採用は見送られた。このため、画像の光量制御はカメラ撮像機構のソフトウェア処理とシャッタースピードの変更でのみ行うものとする。

・事前撮影画像

事前撮影画像は、外出計画に含まれる経路上で、その歩行方向を向いて撮影を行い、同時にその撮影地点と撮影方向を記録する。

1枚1枚撮影時に静止して記録をつけてもよいが、何らかの形で撮影位置を比較的正確に自動推定できる場合は、移動しながらの撮影でもよい。実際、本研究計画では、事前撮影に当たっては経路の一定区間を等速歩行し、それによって各撮影画像の位置を自動推定から得ることもしている。

事前撮影に関しては、日照や天候等様々な要因が実際の歩行時までに変化するため、同一地点同一方向であっても異環境下で撮影した複数枚の画像を用意できることが望ましい。

・参照画像集合

事前撮影画像のうち、外出計画に含まれる経路上のものを参照画像、その集合体を参照画像集合と呼ぶ。

本研究では、歩行揺動に対処するため、画像特徴量には位置・回転・大きさに不変なSIFT特徴量を採用する。このため、参照画像についてはすべて予めSIFT特徴量を求めておく。SIFT特徴量は、実際には1枚の画像につき数十～数百のSIFTキーとして得られる。なお、シーンベースの画像認識ではシーン中の非常に細かい部分の情報は利用しないので、SIFT特徴量算出前に、各参照画像は横320画素縦240画素まで縮小するものとする。

・画像照合

歩行時には、歩行者カメラで撮影を随時行い、システム内の参照画像集合と照合する。システムの都合上、歩行者カメラでの撮影は数秒ごとに連続2枚である。実行時間削減のため、このうちの1枚(入力画像)のみを横320画素縦240画素に縮小した上で、SIFT特徴量を求める。入力画像から得られた各SIFTキーについて、

参照画像集合中の最尤 SIFT キーを求めていく。最も多くの最尤 SIFT キーを含む参照画像を参照画像候補と呼び、その最尤 SIFT キー数をペア数と呼ぶ。また、この最尤 SIFT キーを求める時のキー同士の類似度の算出基準を定めるパラメータを類似度係数と呼ぶ。

ペア数が一定基準を超えた場合、参照画像候補は十分に入力画像に類似していると見なして、その参照画像候補に付与されていた撮影位置と撮影方向情報を出力する。

プログラムとしての実行時には、入力画像の各 SIFT キーに最尤 SIFT キーを求める部分がかつとも計算コストが高い。SIFT キーは 128 次元数値ベクトルで表現されるため、本研究では多次元ベクトル検索の高速化手法の一つである K-D tree 探索を用い、この部分の計算時間を大幅に削減することに成功した。

(倫理面への配慮)

本研究は、その性質から、歩行者カメラで画像を獲得さえすれば実験と解析が可能であり、画像獲得に際して視覚障害者の被験者を用意する必要はない。そのため、データ収集に際して歩行者役には晴眼者を用い、周囲の安全に注意を払って撮影を行った。

また、撮影に際しては、許可が必要な経路においてはしかるべき了解を経路の管理者から得て行った。

撮影した画像については、用途外利用のないよう、細心の注意を払って管理している。なお、画像照合方法の性質上、事前に撮影した画像は画像の形で保管しておく必要はなく、特徴量ベクトル集合、具体的には SIFT キー集合の形で利用され、SIFT キー自体には元画像の表現能力はない。そのため、将来の運用に当たっては、画像データベース用の撮影を行った直後に SIFT キー集合の算出を行って画像そのものは廃棄してしまうことで、画像内に写る人々に対するプライバシー保護も問題ないと考えられる。

C. 研究開発結果

本手法は東芝 dynabook SS RX2 (Core 2 Duo SU9400/ 1.4GHz/ 3MB) 上で実行できるように、OpenCV1.1 と SIFT 公開プログラムを

用いて実装された。

検証実験としては、東京駅八重洲地下中央口付近から地下街を 300m ほど歩き、そこで地上に出てほどなく高島屋に到達、そのあと東京駅八重洲北口近辺まで戻る地上経路約 600m の合計約 900m を外出計画における経路として設定した。経路図を図 II.1-1 に示す。上段赤色の矢印が地下経路に相当し、下段青色の矢印が地上経路に相当する。

同経路の参照画像集合については、2010 年 2 月 20 日に収集した 238 枚と 2010 年 3 月 4 日に収集した 696 枚の合計 934 枚を用いた。実験結果について、ここでは 2010 年 3 月 20 日の午後と夕方の 2 回の歩行時の入力画像に対する位置推定結果を示す。なお、参照画像収集時も入力画像収集時も同一経路を同一方向に歩いているが、歩行速度や歩道内での位置は特に指定しない状態で被験者は歩いたため、これらは均一ではないことに注意されたい。

なお、図 II.1-2 に入力画像収集に用いたシステムの様子を示す。撮影地点は地下街から地上に出た地点である。写真の中で、腹部前面に見える水色の直方体上の部分がカメラ本体、そこから左へ伸びている円筒がレンズである。

図 II.1-3、図 II.1-4 に画像照合結果の様子を示す。左が入力画像、右が検索された参照画像候補である。上下は同じ画像であるが、下段の図中の緑円 1 つが画像中から得られた SIFT キー 1 つを示す。見てわかるように、SIFT キーは数十から数百程度検出される。入力画像中の SIFT キーと参照画像候補中の SIFT キーとが合致した場合、その SIFT キーを赤線で結んでいる。

シーンベースの認識において、シーン中のどこを頼って認識するかは SIFT キー同士のマッチングに従って自動的に決定される。図 II.1-3 では路面の模様、図 II.1-4 では路上の店の看板に主に反応していることが分かる。また、日照条件が異なっていることにも注意されたい。なお、より多くの画像照合の様子を末尾(「H. 付録：画像照合結果の解析」)に示す。



図 II.1-1 実証実験経路



図 II.1-2 画像取得に用いたシステム

実際の歩行支援システム稼働時には、他のサブシステムとの負荷バランスの都合もあり、およそ8～10秒に一度、本手法を実行することになっているが、ここでは検証のため、3月20日の歩行時に経路上で得られた全ての画像に対する認識結果を表 II.1-1, II.1-2 に示す。表 II.1-1 は当日昼の歩行、表 II.1-2 は当日夕方の

歩行に対する実験結果である。表では、類似度係数の変更が認識率にどのような結果を及ぼすかを示している。類似度係数を1.0に近づけるほど、ペア数が増加していく。参照画像候補の検定を行うためのペア数閾値を2～7にした場合の、返答数と返答率も併せて示す。問い合わせに対して、表 II.1-1 では平均返答率16.8%、表 II.1-2 では35.9%だった。これは、表 II.1-2 の夕方の状況のほうが、同日昼よりも、2月20日および3月4日の天候状況に近かったためと考えられる。

平均誤答率を0.1%以下にした場合の類似度係数と有効返答率平均値の関係を表 II.1-3 に示す。

表から分かる通り、誤差率0.1%以下でも類似度係数を0.9にまで上げれば有効返答率を向上させることができるので、類似度係数については0.9を採用する。その条件下で、ペア数の閾値に対する返答率平均値と誤答率平均値の変化を調べた結果が表 II.1-4 である。ペア数閾値を下げ過ぎると、誤答が増え始めるので、ここではペア数閾値を6とする。

こうして得られた類似度係数0.9、ペア数閾値6をシステムパラメータとして組み込むことで、入力画像から誤認識を避けつつ撮影位置と撮影方向の推定を高い確率で求めることができるようになった。



図 II.1-3 画像照合の様子その1

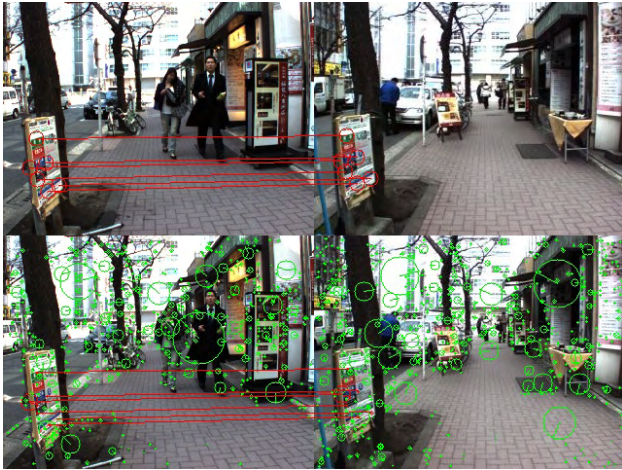


図 II.1-4 画像照合の様子その 2

表 II.1-1 3月20日昼の歩行に対する結果

類似度係数	0.6	0.7	0.8	0.9
ペア数平均値	1.21	2.26	3.85	6.91
ペア数閾値	2	2	4	7
入力画像枚数	1252	1252	1252	1252
有効返答数	126	354	304	277
有効返答率	10.1%	28.3%	24.3%	22.1%
誤答数	1	1	1	0
誤答率	0.08%	0.08%	0.08%	0.00%

表 II.1-2 3月20日夕方の歩行に対する結果

類似度係数	0.6	0.7	0.8	0.9
ペア数平均値	2.79	4.64	7.12	11.01
ペア数閾値	2	4	5	6
入力画像枚数	1269	1269	1269	1269
有効返答数	481	448	575	729
有効返答率	37.9%	35.3%	45.3%	57.4%
誤答数	1	0	0	0
誤答率	0.08%	0.00%	0.00%	0.00%

表 II.1-3 類似度係数に対する有効返答率

類似度係数	0.6	0.7	0.8	0.9
有効返答率平均値	24.0%	22.2%	30.8%	33.8%

表 II.1-4 類似度係数 0.9 時の結果

ペア数閾値	3/20(昼)		3/20(夕方)	
	返答率平均値	誤答率平均値	返答率平均値	誤答率平均値
5	35.7%	0.56%	64.9%	0.16%
6	26.3%	0.12%	57.4%	0.00%
7	22.1%	0.00%	51.4%	0.00%
8	18.7%	0.00%	45.5%	0.00%

本システムをある地上歩行(2月20日午後)時に採取した画像に適用した結果を図 II.1-5 に示す。横軸は画像番号である。撮影は歩行開始から約1秒間隔で2枚ずつ撮影したので、概ね9分程度の歩行に相当する。縦軸は参照画像番号を示している。参照画像集合は図 II.1-1 に示した経路全てを2周回って収集したもので、最初の0~237番が1週目(2月20日撮影)、238番から933番までが2週目(3月4日撮影)に相当する。2週の参照画像取得いずれも普通歩行で実験を行ったが、撮影間隔が2倍程度違うことに注意されたい。入力画像も普通歩行であるので、グラフはほぼ直線状になるのが理想である。ただし、経路には信号等があるため、その部分はグラフはほぼ水平になる。また、歩行者カメラの前に別の歩行者等がいる間は、画像参照にあまり成功しないため、その部分是对応付けができないことになる。その一方で、ご対応があると、この直線状の配列から大きくはずれた点が誤対応としてグラフ上に現れる。グラフからも、誤対応がほぼ存在しないことがわかる。

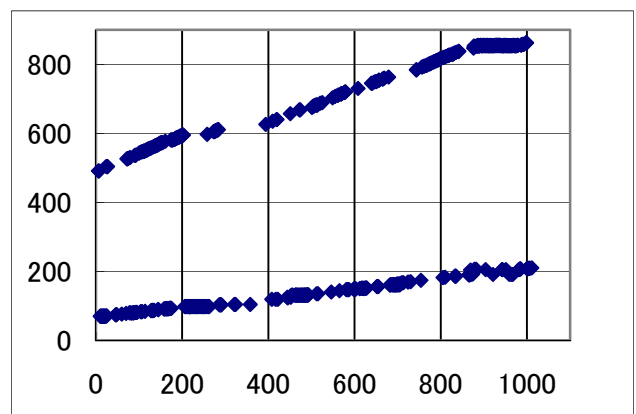


図 II.1-5 2月20日午後歩行に対する実験結果

一方で、表 II.1-4、図 II.1-4 から分かるように、返答率は状況にもよるが問い合わせに対して、1～4割程度であり、残りの9～6割においては、位置推定不能という解がシステムとしては返されることになる。誤認識率を許容すれば返答率を向上させることは可能であるが、視覚障害者歩行支援システム全体としてシーンベース画像認識プログラムに要求されることは返答率の向上ではなく、信頼できる推定値の提供である。そのため、誤答率を最小にするほうをパラメータ決定時には重視している。

参考までに、現行の全体システムでは本プログラムは8～10秒に1回実行されることになっている。仮に10秒に1回だとすると、0.1%の誤等率は10,000秒間（2時間以上）に対して1回程度だけ誤答することに相当する。

D. 考察

本研究計画の分担研究では、シーンベース画像認識に基づく撮影位置・方向推定を、視覚障害者支援システムの一部として基本的に実現することができた。

ただし、シーンベース画像認識は、撮影環境に大きく影響を受ける。以下、個別項目について考察をしていく。

・カメラの取り付け方向

本研究計画で想定する歩行経路は、地下街から階段、大通りから細い歩道まで多岐に渡る。カメラの取り付け方向としては、進行方向前方の他に、進行方向に対して水平90°右ないし左方向、進行方向のまま仰角ないし俯角をつけることが検討された。

進行方向前方にカメラを向けた場合、概ね進行方向に沿って歩道が広がっていることになる。特に遠方の風景の見かけは距離に対して概ね反比例の大きさで見えることになり、現在利用している SIFT 特徴量は大きさ不変なので、参照画像候補は正しい撮影位置に対して多少前後する可能性がある。この推定精度を高めるためには、画像間の見えているシーンの大きさの違いを計測し補正することが考えられる。

一方、それ以外の方向にカメラを向けた場合、

いずれも相対的に画像中における近隣物体像が占める割合が増加することになる。例えば道端にある店へ入るドアの位置を確認しやすくなるなど、位置推定精度の向上に寄与する可能性があるが、実際に予備実験をしてみると、被撮影物体が近い場合には画像照合が困難になる傾向が顕著にみられた。これは、画像照合の元となる SIFT キーが、被撮影物体に対する撮影方向の変化が±20°程度を超えると、たとえ同じ物体を撮影していたとしても異なる SIFT キー表現になってしまうことが主たる原因と推察される。そのため、今後も歩行者カメラは進行方向前面に向けるように設置していく予定である。

・カメラ撮像機構

画像解像度はそれほどなくとも画像照合の性能低下が起きないことが予備実験から確認されている。そのため、今後はグローバルシャッターの VGA 程度の解像度のカメラを想定し、レンズと合わせて小型化を目指す。

入力画像に関しては、現行システムはキャプチャ時に縦1024画素横768画素で画像を取り込み、そのあと縦320画素横240画素への縮小をソフトウェアで行っているため、数百ms程度実行時間がかかっている。この部分は画像取込時からサイズを指定することで削減可能である。

・レンズ

屋内屋外の両方に対応するためには、カメラ撮像機構と合わせてオートアイリスが可能であるか、安価で小型なカメラ系を2つ用意して、一方を屋内の光量、もう一方を屋外の光量に調整することが考えられる。現状では、晴天等の場合、地下や屋内との調整がしきれずに画像情報の欠損が見受けられた。

・画像照合

本研究計画では SIFT 特徴量を用いたが、他モジュールが SURF 特徴量を用いていること、予備実験から SIFT と SURF でそれほど性能差がないと見られていること、SURF のほうが必要な計算コストが小さいことから、今後は SURF に移行していくほうがよいと考えられる。

また、現在のシステムの懸案事項として、

1000 枚規模の参照画像集合の読み込み時に 1.5GB 程度のメモリを一時的に使用し、かつ実行時に 500MB 程度のメモリが必要であることが挙げられる。解析により、参照画像集合の中で実際に SIFT キーとして有効なのは実験的には半分以下であろうとの予想が立てられている。起動時の一時的なワークエリアの確保も不要にできるとの技術的検討ができていたので、この部分のメモリ使用量は大幅に削減できることが見込まれている。

その一方で、シーンベースの画像照合の性能向上のためには、同一地点で撮影環境の異なる複数の画像を用意することが有効であるという予備実験結果も得られている。K-D tree 探索は参照画像枚数の増加に対して検索時間の増加率が小さくすむという利点があるため、今後の展開としては、上記のメモリ削減を行った上で、参照画像集合の枚数を増加させることを検討していくべきである。

・参照画像集合

本研究計画実施に当たっては、ローカルに参照画像集合を用意し、予備実験等を行う必要があった。そのために収集した画像群のうち、撮影の位置・方向情報を付与したデータ取得履歴を示す。いずれも東京駅八重洲口から高島屋に至る一帯で撮影したものである。

2009/11/26	519 枚
2010/01/28	1,774 枚
2010/03/04	3,038 枚

また、位置・方向情報を付与していないが、予備実験用のデータとして、2009 年 11 月から 2010 年 4 月に至る 5 ヶ月間で 50 本以上の経路に沿ったビデオ撮影を行った。なお、経路は本報告での「C. 研究開発成果」(図 II.1-1)で述べた東京駅から高島屋を経由して東京駅に戻るものである。

E. 結論

本研究計画では、シーンベース画像認識に基づいて、歩行者カメラの撮影位置・方向推定方法を提案し、システムを実現した。現行のノートパソコンレベルのハードウェアパフォーマンスで、画像獲得から位置方向推定までを 2 秒以

下でほぼ実行可能である。

開発期間が短かったこともあり、システムの完成度はまだプロトタイプレベルの部分があるので、今後の展開のためには「D. 考察」で述べたようにプログラムを洗練させていくことが必要である。これらはいずれも予備実験等を通して一定の目処がついているため、いずれも実現に対しては特に問題ないと考える。

F. 研究発表

1. 学会発表

- ・査読付き国際会議投稿中 1 件 (採否通知 4 月 7 日予定)
- ・電子情報通信学会パターン認識・メディア理解研究会 (2010 年 5 月発表予定)

H. 付録：画像照合結果の解析

以下は 3 月 20 日の実験の様子を示す。実験歩行の順に Data set-3 ~ Data set-6 と呼ぶ。この順に時間を空けて撮影したので、左側の入力画像が夕方に向かって徐々に暗くなっていく様子にも注目されたい。なお、図 II.1-3,II-1-4 は Dataset-3 での実行結果である。

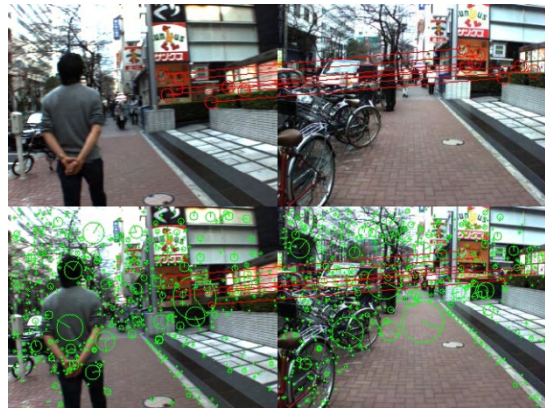
Dataset-3/#1

高島屋後、北西に向かって伸びるさくら通り上。さくら通りは狭くかつビルが高いため、一日を通じてほとんど日が差さない。本結果では左側の街路樹の枝形状も利用している。このような自然物の特徴は、春以降は残念ながらほぼ利用不可能になると考えられるが、この場合は他にも看板等があるので対応可能である。



Data set-4/#1

同じくさくら通り上。看板が利用されている。すぐ上にあるサンクスの看板は露光の関係で SIFT キーレベルでは同一物体とは判定されていない。



Data set-4/#2

同じくさくら通り上。自動販売機のパターンが認識に利用されている。画像左側や前方の歩行者にも SIFT キーは数多く発生しているが、同じものがないため照合には寄与していない。

