

(71) 文字認識とSFMによる携帯端末ユーザの 自己位置推定

安室 喜弘¹・山森 康生²・廣瀬 詢⁴・藤里 和樹⁴・檀 寛成³

¹正会員 ²学生 ³非会員 関西大学環境都市工学部都市システム工学科

⁴学生 関西大学理工学研究科環境都市工学専攻都市システム工学分野

(〒564-8680 大阪府吹田市山手町3丁目3番35号)

E-mail: yasumuro@kansai-u.ac.jp

近年、スマートフォンの普及により位置情報を利用したサービスが普及しているが、屋内や地下など、GPS信号が届かない場所では、自分の位置情報を得る手段が確立されていないのが現状である。本研究では、位置情報を知る手掛かりとして、地下街の環境中に散見される文字情報に着目し、携帯端末のカメラにより自己位置を推定する方法を提案する。現地で収集した写真から文字情報を自動認識するとともに写真測量の原理で3次元復元しておくことで、シーンに含まれる文字情報をランドマークとする。PnP問題を単純化し計算コストを抑えた実装で、2m程度の分解能で地下街で自己位置を検出できることを確認した。

Key Words : indoor location, character recognition, SFM (structure from motion)

1. はじめに

近年、携帯端末の普及により地図情報を利用したサービスが身近なものとなっており、導入コストのかからないアプリケーションでさえ、ナビゲーション、店舗案内など動的で幅広いサービスを提供している。しかし、屋内や地下など、GPS信号が届かない場所に入った途端に、これらのサービスは機能しなくなる。屋内での位置推定手法には、ビーコン^{1,2)}やWi-Fi電波や地磁気等の電磁波を利用する方法³⁾のほか、端末のカメラ画像の自然特徴点を利用した手法がある^{4,5)}。ビーコンなどの付帯設備にはコスト面で難点があり、電磁波環境は時間帯や場所によって不安定な場合がある。画像を使う方法は、自然特徴点が多いほど頑健であるが、計算コストも多大になる。本研究では、画像情報に基づいた手法を基本とし、都市部で発達している地下街や交通機関などの環境において、文字情報が豊富であることに注目し、経路案内に堪える程度の速さと正確さを併せ持つ位置推定手法を提案する。

2. 提案手法

一般に、画像の2次元座標と空間の3次元座標の対応か

らカメラの位置・姿勢を求める問題はPnP (Perspective-n-Point) 問題として知られ、先行研究では、精度を担保するために多くの自然特徴点について処理が行われる。ここでは、厳密なカメラの位置や姿勢ではなく、ユーザの立ち位置を求める目的から、次のように単純化する。まず、サービス対象とする地下空間において視点と方向を変えながら撮影した写真を収集する。収集した画像に対してSFM (Structure From Motion) を適用して対象空間の3次元モデルを作成する。次に、収集した各画像において、文字認識処理を行い、識別された文字について画像座標と、SFMによる3次元座標との対応づけを行う。新たにユーザ端末で撮影された画像に対しても、同様に文字情報を抽出し、事前に登録された情報を参照して、対応する3次元座標を取得する。1つの画像内で複数の文字について、画像上の2次元位置と、3次元モデルの3次元座標が得られる。図-1に示すように、画像上で識別された異なる文字 (列) i, j の2次元座標とカメラの画角との関係から、カメラの焦点 O_c から各文字の方向を指す3次元ベクトルを求める。これら複数の文字の方向ベクトルについて互いの内積 θ_{ij} を計算しておく。一方、識別された文字 (列) i, j のSFMによる3次元座標についても、ユーザ位置の候補となる座標 O_u から各文字を指す方向ベクトルを求め、同様に互いの内積 θ_{ij} を計算す

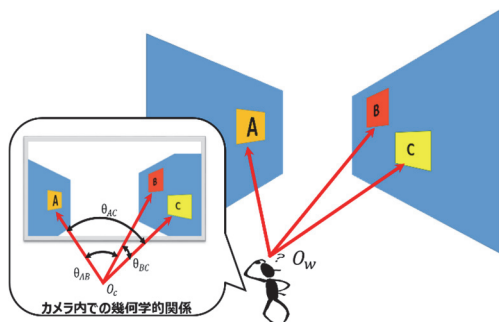


図-1 基本アイデア

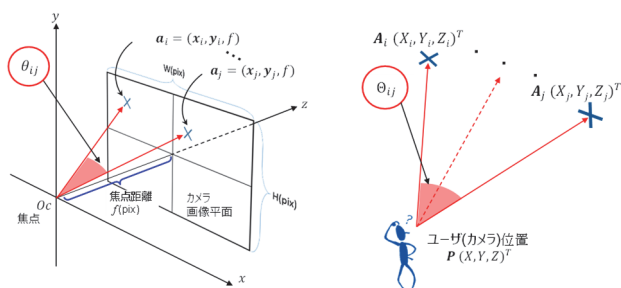


図-2 文字 (列) i, j の幾何的關係

る. 2つの内積の値 θ_{ij} と $\theta_{ij}(P)$ の二乗誤差をとり, 画像内の異なる文字 (列) の全組合せについての総和を最小とするようなユーザの立ち位置 P を解とする. (式(1)参照)

$$\operatorname{argmin}_P J(P) = \sum_{i,j \in \text{文字列集合}, i \neq j} (\cos \theta_{ij} - \cos \theta_{ij}(P))^2 \quad (1)$$

3. 実験

(1) 実験概要

人通りが多く, 多種多様な店を擁する場所として, 梅田地下街 (大阪駅前第3ビルB2F) の一角を対象として実験を行った. ウェアラブルカメラ (Panasonic製HX-A500: 3840×2160画素, 水平画角136°) により通路を12m程度通行しながら撮影した画像34枚をVisual SFM⁹⁾により3次元復元を行い, さらにメッシュ化して3次元モデルを生成した. 図-2に生成したモデルを示す. 文字認識にCloudVision APIを用い, 文字 (列) の中心の画像座標を記録した. 文字 (列) 座標と3次元モデルでの座標との対応付けを行った上で, カメラの焦点距離と画角, 画素数の関係から, カメラのローカル座標系 (3次元) における文字の方向ベクトルを求めた. ここでは図-3に示すように, 「和牛」「GAME」「エレベータ」「地下鉄」の文字列を用いた場合, 通路上の3.7m×12.0mの領域を対象としてユーザの位置を推定する. ユーザの立ち位置の領域を9×16グリッドに分け, グリッドの各点で内積値の二乗誤差値を算出する. なお, 歩行者の身長を考慮して, 床面から1.7mに相当する高さで固定した. 算出された誤差の分布を図-4に示す. 1枚の写真入力から, 進行方向については0.75~1.50m (1~2グリッド) 程度, 横方向



図-3 現地の3次元復元モデルとユーザ位置検出対象領域

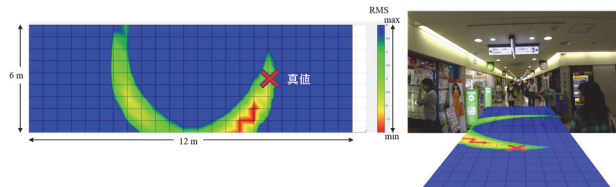


図-4 ユーザ位置推定結果例

については, 1.2~1.8m (4~6グリッド) 程度の位置推定分解能が得られた. 誤差分散に方向性が見られるのは, 写真を撮影した視線方向と, 検出された文字列の方向に依るものと考えられるが, 通路での位置として, どのブロック, どの店舗の付近にいるかなど, 行動の記録や, 経路の誘導などに有効な位置情報であると考えられる.

4. おわりに

本研究では, SFMと文字情報を用いたユーザ携帯端末位置推定手法を提案し実験により有効性を確認した. 今後は, 観測される文字情報の分布や指向性について分析を進める予定である.

謝辞: 本研究の一部は独立行政法人日本学術振興会 科学研究費補助金 (15H02983, 16H05000) および関西大学先端科学技術推進機構研究グループの助成による.

参考文献

- 1) 岩見泰周, 鬼木明日香, 佐藤健哉: “複数iBeaconを利用した位置精度向上の検討”, 情報科学技術フォーラム講演論文集14(4), pp.309-310, 2015.
- 2) 藤田迪, 佐々木威, 河口信夫: “地下鉄構内の制約条件を利用した無線LAN位置推定手法”, 情処報告 (UBI) 66, pp.153-157, 2008.
- 3) 山崎俊太郎, 持丸正明, 金出武雄: “一人称ビジョンシステムのための自己位置推定法”, 信学技報, pp.73-78, 2010.
- 4) 佐藤智和, 池田聖: “景観画像データベースを用いたユビキタス端末の自己位置推定に関する研究”, SCOPE 第6回成果発表, pp.20-21, 2010.
- 5) Wu C.: Toward Linear-time Incremental Structure from Motion, 3DV 2013, pp.127-134, 2013.