

K-006

位置情報付き俯瞰画像を用いたモバイルカメラの位置と方位の推定 A Mobile Camera Localization by Using an Overhead-View Image with Positional Information

鳥屋 剛毅† 北原 格† 大田 友一†

Hisatoshi Toriya Itaru Kitahara Yuichi Ohta

1. はじめに

GPS (Global Positioning System) を搭載した携帯電話やモバイルカメラの普及を背景として、GPS や電子コンパスで獲得されるユーザの位置・方位情報を用いた街角案内システムが実用化されている。一般的な GPS の位置情報には、数メートル程度の誤差が含まれ [1]、都市部などの複雑に入り組んだ道や、GPS 衛星の電波がマルチパスを起こすビル群などでは、さらに誤差が増え、その結果誤った情報指示がなされてしまう問題が知られている。方位情報に関しては、自動車に搭載された GPS では時系列測位により進行方向を推定しているが、モバイル端末に搭載された GPS では、ユーザは立ち止まったり、ゆっくり歩いたりするため、時系列観測による方位推定が困難である。電子コンパスの使用も考えられるが、電子コンパスの方位情報には 0~5 度ほどの誤差が含まれ [2]、環境によっては周囲の電子機器などで発生した磁場により誤差が生じる。

一方で、近年、Google Earth [3] を代表として、人工衛星や航空機から地上を撮影した俯瞰画像の入手が容易になりつつある。地理情報システム (GIS: Geographic Information System) の俯瞰画像データは、各画素に位置情報を含んでいる。そのような位置情報を持った俯瞰画像とモバイルカメラ画像中の画素とを対応付けることができれば、より正確にモバイルカメラの位置・方位を求めることが可能となり、図 1 に示すような、自己位置・方位の提示システムなどへの応用が期待できる。

本論文では、位置情報付き俯瞰画像とモバイルカメラで撮影した画像を対応点探索することにより、モバイルカメラの位置・方位を画像ベースで推定する手法を提案する。

2. 関連研究

モバイルカメラの位置・姿勢を推定する手法は、センサベースの方法と画像ベースの方法に大きく分けられる。前者の例としては GPS や電子コンパスなどを利用する手法が考えられるが、一般的に用いられる GPS は測位誤差が大きく [1]、電子コンパスも誤差を含むので [2]、高精度な推定が困難である。後者の例としては、ARToolKit [4] などのようにマーカーを用いる手法がある。この手法は、精度の良い推定が可能であるという利点があるが、画像中に常にマーカーが写っている必要があることや、マーカー設置により撮影空間の景観が損なわれるといった問題が知られている。

環境カメラ設置台数の増加を背景に、環境カメラを併用したモバイルカメラの位置姿勢を推定する研究が行われている。

水流ら [5] は、屋内環境においてモバイルカメラとステレオカメラ (環境カメラ) を併用し、モバイルカメラの位

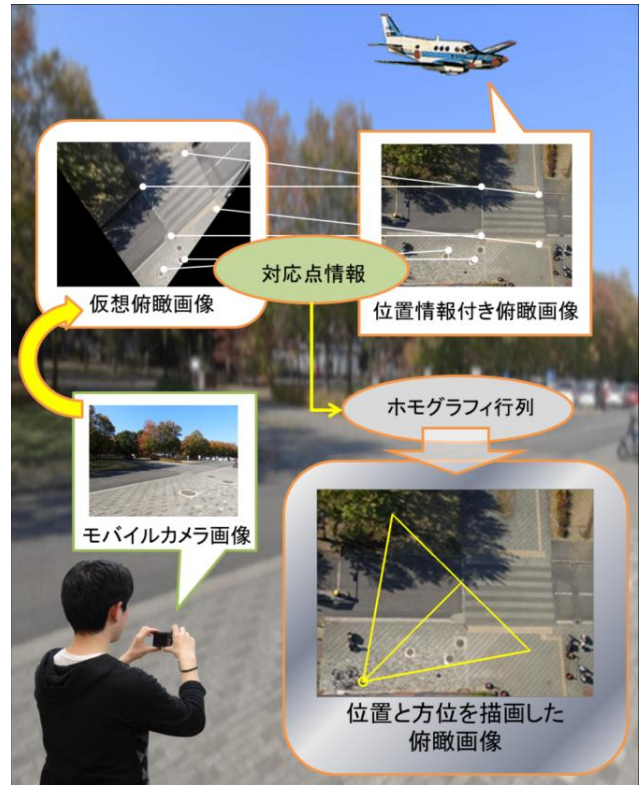
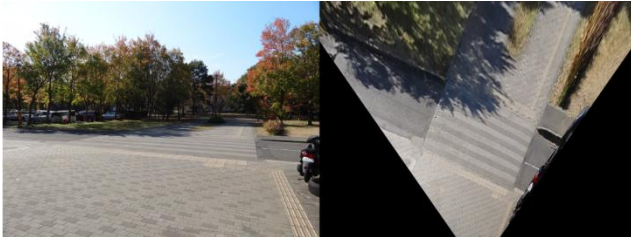


図 1: 目標とする位置と方位の提示システム

置・姿勢の推定を実現している。また、野田ら [6] は車載カメラと、上空から航空機などで真下を撮影した俯瞰画像を併用し、自動車の位置推定を行う手法を提案している。車載カメラによって撮影した画像から手動で道路領域を抽出し、仮想的に上空から見た俯瞰視点へと変換し、空撮画像と対応点探索をすることによって求まるホモグラフィ行列より自車位置の推定を行っている。

本研究では、環境カメラの画像として、高所より真下を撮影した俯瞰画像に注目した。近年、Google 社の提供する Google Earth [3] や国土地理院の提供する GIS の空中写真閲覧サービスなど、俯瞰画像の入手が容易になりつつある。俯瞰画像は、上空から撮影するため、撮影回数が少なくても広範囲の情報を得られるという利点を有する。本研究では、これらの俯瞰画像と人間が持ち運びできる、コンパクトデジタルカメラ画像の対応点探索を行うことにより、モバイルカメラの位置・方位を正確に推定する手法を提案する。

† 筑波大学大学院システム情報工学研究科



モバイルカメラ画像 仮想俯瞰画像

図 2: 仮想俯瞰画像生成の例

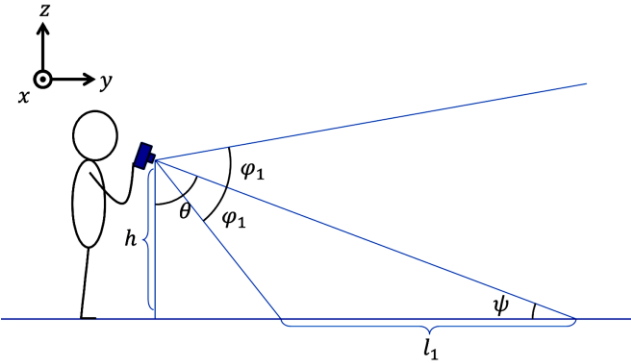


図 3: x 軸方向から見たモバイルカメラと地面領域

3. 俯瞰画像を用いたモバイルカメラの位置と方位の推定手法

ユーザは、モバイルカメラを用いて目の前の風景を撮影する。カメラの内部パラメータと撮影時のチルト角に基づいて、撮影した画像を俯瞰視点に射影変換し、仮想俯瞰画像を生成する。仮想俯瞰画像と高所から実際に撮影した俯瞰画像において SIFT [7]を用いた対応点探索処理を行った結果から、両画像間のホモグラフィ行列を算出する。そのホモグラフィ行列を用いて、モバイルカメラの位置と方位を推定する。

モバイルカメラ画像と俯瞰画像間の対応点探索を行う際、画像中での回転にロバストであるという SIFT の特長を活かし、本手法では、モバイルカメラからの見え方を俯瞰視点からの見え方に変換して対応点探索を行う。

野田ら [6]の研究で用いられている車載ステレオカメラとは異なり、手持ちのモバイルカメラは位置・姿勢が不安定であり、必ずしも位置・方位推定の処理に適した写真が得られるとは限らない。そこで、本手法では、ユーザに簡単な撮影補助を依頼することにより、処理精度を向上させる。具体的には、画面を上下に二分割する線を引き、この線よりも上に地平線が水平に写るよう撮影をしてもらうことにより、

- モバイルカメラの光軸が地面と交わる
- モバイルカメラのロール角は無視できる程度に小さい

という仮定を満たす画像の獲得が可能となる。これらの仮定に基づき、次節で述べる手法により、仮想俯瞰画像を安定して生成することが可能になる。

4. 仮想俯瞰画像生成手法

モバイルカメラ画像から、図 2 に示す仮想俯瞰画像を生成する手法について述べる。図 3 はユーザが撮影している様子を真横から見た図であり、図 4 は真上からの図である。撮影前にあらかじめ計測しておく必要があるものは、図 3

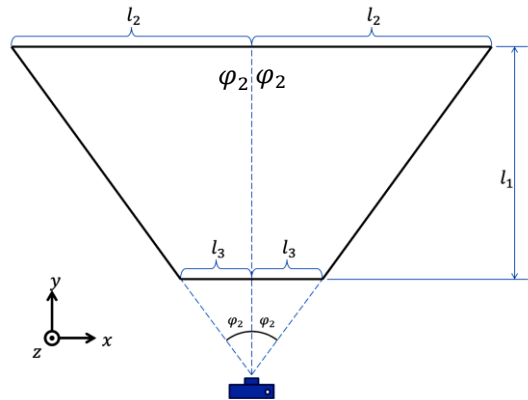


図 4: z 軸方向から見たモバイルカメラと地面領域

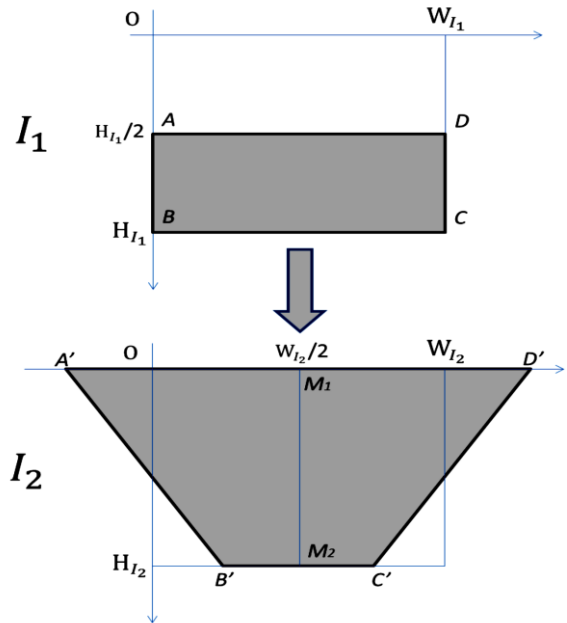


図 5: カメラ画像における地面領域 4 点の変換先

に示す垂直面角 $2\varphi_1$ と、図 4 に示す水平面角 $2\varphi_2$ である。撮影の度に計測する必要があるのは、図 3 に示すチルト角 ψ であり、垂直面角 $2\varphi_1$ と水平面角 $2\varphi_2$ はカメラパラメータから、チルト角 ψ は重力センサなどを用いれば取得が可能である。モバイルカメラ画像を仮想俯瞰画像へ射影変換するホモグラフィ行列 H_1 を求めるためには、変換前の画像と変換後の画像間において、最低 4 点の対応が必要である [8]。本手法では、その対応点の座標を、上述した定数 φ_1 , φ_2 と、変数 ψ から算出する。図 3 に示すように h , $\theta = \pi/2 - \psi$ をとると、

$$l_1 = h \tan \theta - h \tan(\theta - \varphi_1) \quad (1)$$

となり、また、図 4 に示すように l_2 , l_3 をとれば

$$l_2 = h \tan \theta \tan \varphi_2 \quad (2)$$

$$l_3 = h \tan(\theta - \varphi_1) \tan \varphi_2 \quad (3)$$

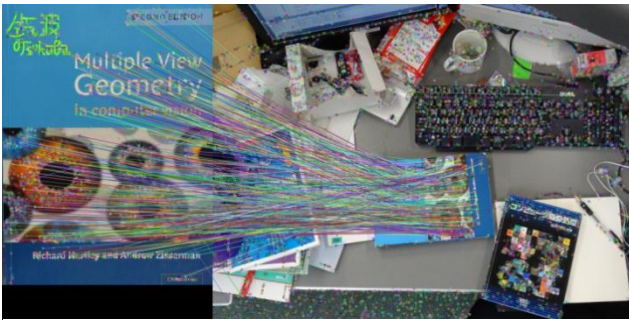


図 6 : SIFT による対応点探索の例

図 7 : H_2 で射影変換された入力画像と、参照画像

と表せる。

入力画像 I_1 と出力画像 I_2 のサイズがそれぞれ $H_1 \times W_1$, $H_2 \times W_2$ であるとする。図 5 は, I_1 の地面領域 $ABCD$ が I_2 においてどう変換されるかを表したものである。図 3 の l_1 が H_2 となるように変換されていることがわかる。辺 $A'D'$, $B'C'$ の中点をそれぞれ M_1 , M_2 とすると,

$$H_2: \overline{A'M_1} = l_1: l_2 \quad (4)$$

$$H_2: \overline{B'M_2} = l_1: l_3 \quad (5)$$

が成り立つので,

$$\overline{A'M_1} = H_2 \times \frac{l_2}{l_1} \quad (6)$$

$$\overline{B'M_2} = H_2 \times \frac{l_3}{l_1} \quad (7)$$

つまり

$$\overline{A'M_1} = H_2 \times \frac{\tan \theta \tan \varphi_2}{\tan \theta - \tan(\theta - \varphi_1)} \quad (8)$$

$$\overline{B'M_2} = H_2 \times \frac{\tan(\theta - \varphi_1) \tan \varphi_2}{\tan \theta - \tan(\theta - \varphi_1)} \quad (9)$$

となる。これらの式から求まる 4 点の変換後の座標から、ホモグラフィ行列 H_1 を算出し、それを用いてモバイルカメラ画像を仮想俯瞰画像に変換する。

5. 仮想俯瞰画像と俯瞰画像間の対応点探索

本節では、仮想俯瞰画像と俯瞰画像との対応点探索手法と、その結果を用いて仮想俯瞰画像を俯瞰画像に射影変換するホモグラフィ行列 H_2 を推定する手法について述べる。

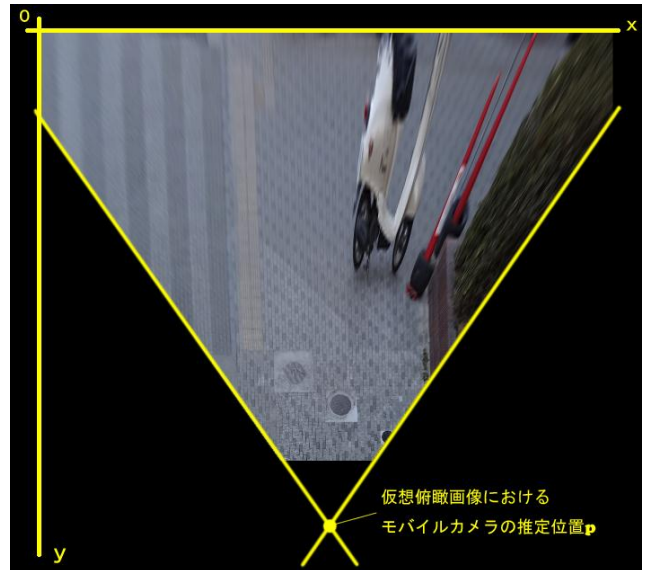


図 8 : 仮想俯瞰画像におけるモバイルカメラの推定位置

5.1. 対応点探索の手法

対応点探索には、SIFT を用いる。SIFT は画像中から特徴点を検出し、それらの各特徴点の特徴量を記述するアルゴリズムである [9]。SIFT によって記述された特徴量は、回転、照明変動、スケール変化にロバストであるという特長を有する。モバイルカメラ画像を仮想俯瞰画像に変換することにより、回転とスケールだけが異なる画像間での対応点探索処理となるため、SIFT の特性が効果的に活用される。

5.2. 対応点情報を用いたホモグラフィ行列の推定

SIFT による対応点探索によって得られた対応点の情報から、RANSAC [10] を用いたロバスト推定を行い、仮想俯瞰画像を俯瞰画像に射影変換するホモグラフィ行列 H_2 を算出する。図 6 では、左が入力画像で右が参照画像であり、SIFT によって得られた対応点を線で繋いである。明らかな誤対応が数点存在するが、ロバスト推定によりそれらを除去した結果、図 7 に示すように、適切な H_2 を求めることができる。

6. 位置と方位の推定

4 節と 5 節で、モバイルカメラ画像を仮想俯瞰画像に射影変換するホモグラフィ行列 H_1 と、仮想俯瞰画像を俯瞰画像に射影変換するホモグラフィ行列 H_2 を算出する手法を述べた。これら 2 つの行列の積

$$H = H_2 H_1 \quad (10)$$

は、モバイルカメラ画像を俯瞰画像へ射影変換するホモグラフィ行列となる。本節では、これらのホモグラフィ行列を用いて、モバイルカメラの位置・方位を推定する手法を述べる。

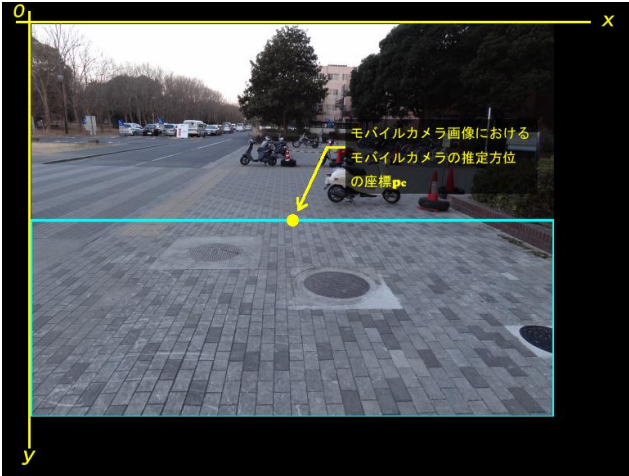


図 9: モバイルカメラ画像におけるモバイルカメラの推定方位



図 11: 実験場所の俯瞰画像 (解像度は 25[cm/pixel])

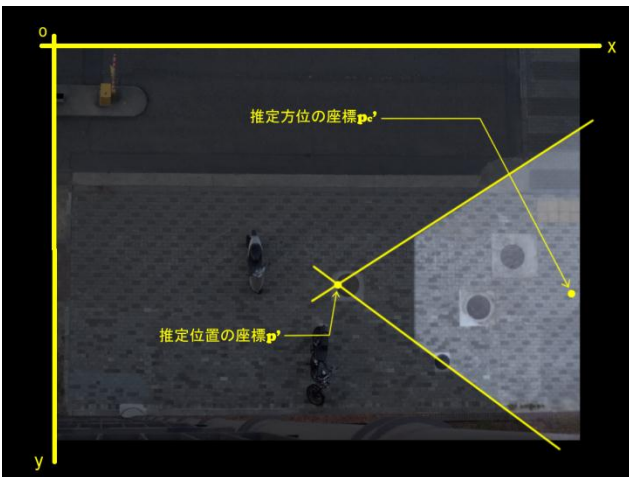


図 10: モバイルカメラの推定位置と方位

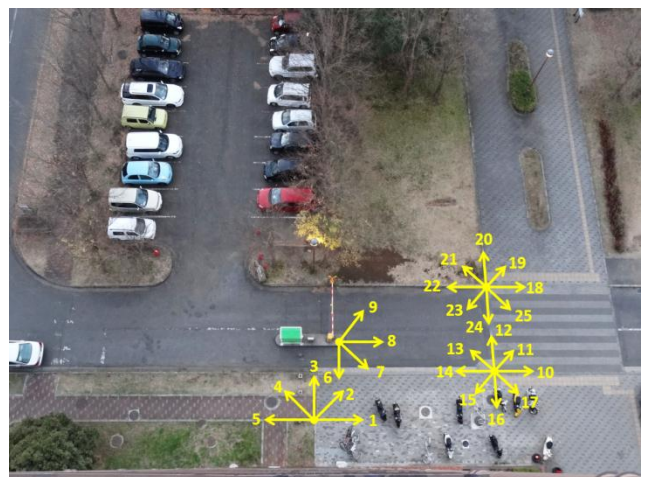


図 12: 実験に用いた俯瞰画像と撮影地点・方向・画像番号

6.1. モバイルカメラの位置推定手法

図 8 に示すように、仮想俯瞰画像におけるモバイルカメラの位置 \mathbf{p} は、二つの斜辺の交点として推定される。具体的には、 \mathbf{p} の座標を同次行列で

$$\hat{\mathbf{p}} = \begin{pmatrix} x \\ y \\ 1 \end{pmatrix} \quad (11)$$

とすれば、これを H_2 で変換した

$$\hat{\mathbf{p}}' = \begin{pmatrix} x' \\ y' \\ 1 \end{pmatrix} = H_2 \hat{\mathbf{p}} \quad (12)$$

は、俯瞰画像におけるモバイルカメラの推定位置の座標となる。仮想俯瞰画像の各画素には、位置情報が対応付けられているため、この処理によって、座標 \mathbf{p}' の画素に対応する位置情報より、モバイルカメラの位置を獲得することができる。

6.2. モバイルカメラの方位推定手法

図 9 に、モバイルカメラ画像の画像中心をマークした図を示す。撮影技法 (フレーミング) によっては、必ずしもユーザの撮影したい対象が画像の中心に写っているとは限

らないが、本研究で言うカメラの方位とは、画像中心の方向であるとする。モバイルカメラ画像の画像中心の座標を

$$\hat{\mathbf{p}}_c = \begin{pmatrix} x_c \\ y_c \\ 1 \end{pmatrix} \quad (13)$$

とすれば、式(10)の H を用いて射影変換し

$$\hat{\mathbf{p}}_c' = \begin{pmatrix} x_c' \\ y_c' \\ 1 \end{pmatrix} = H \hat{\mathbf{p}}_c \quad (14)$$

を求めることで、俯瞰画像における画像中心の座標 \mathbf{p}_c' を求めることができる。図 10 は俯瞰画像にモバイルカメラと画像中心の推定位置をプロットしたものである。 \mathbf{p}' から \mathbf{p}_c' への方向が推定方位である。

7. 実験

2012 年 1 月 18 日午前 7 時、筑波大学第三エリアの工学系 F 棟周辺で位置と方位の推定実験を行った。撮影には電子コンパス搭載のモバイルカメラである、SONY Cyber-shot DSC-HX7V を使用した。入力 of モバイルカメラ画像の画素数は VGA (640 × 480) のものを用いた。

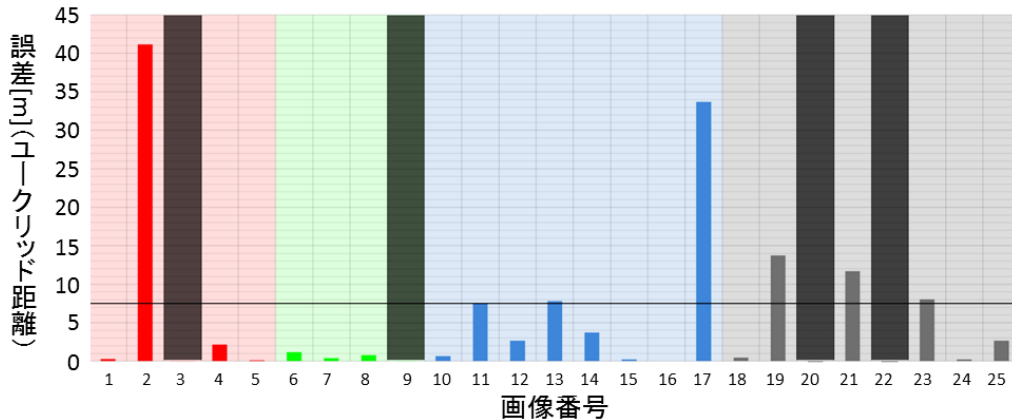


図 13 : 位置推定の誤差のグラフ (ユークリッド距離)

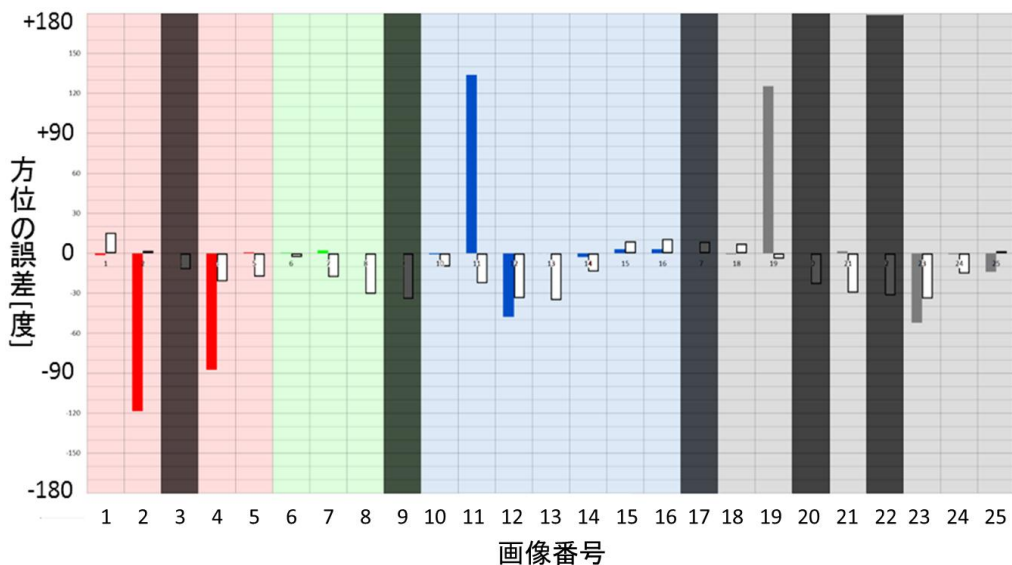


図 14 : 方位推定の誤差のグラフ

7.1. 実験で使用する位置情報付き俯瞰画像

位置情報付き俯瞰画像としては図 11 に示すような、GIS より入手できる俯瞰画像 [11]などが考えられるが、現在一般に入手可能な俯瞰画像は空間解像度が 10~40[cm/pixel]程度であり、対応点探索を行うには解像度が低いといった問題があるため、この実験では図 12 に示すような、2011 年 12 月 11 日午後 4 時に工学系 F 棟の 12 階 (地上約 42[m])よりほぼ真下を見下ろした画像を俯瞰画像として用いることとする。測量を行った結果、空間解像度はおよそ 4.2[cm/pixel]であった。

7.2. 位置と方位の推定実験

図 12 に撮影位置と撮影方位、その方位で取られた画像の番号を示す。4 節の図 5 で定義したモバイルカメラのチルト角 ψ は 10 度にし、各地点の各方向を撮影した。

7.3. 結果と考察

7.3.1. 位置の推定結果

図 13 に、撮影位置と推定位置の誤差をユークリッド距離で表現したグラフを示す。撮影位置によってバーと領域を色分けした。背景が濃い灰色の部分は、対応点の探索に失敗し値が出せなかったことを示している。また、GPS を

用いた位置推定との比較のために一般的な GPS を用いた場合の平均誤差 [1]である 7.5[m]に線が引いてある。

7.3.2. 方位の推定結果

図 14 に、モバイルカメラ画像と俯瞰画像から手動で求めた方位の真値と、提案手法で求めた方位との誤差、それに加え電子コンパスで計測した方位との誤差をグラフにしたものを示す。真値と提案手法で求めた値との誤差を各画像番号の左側に赤・緑・青・灰色のバーで、電子コンパスで計測した値との誤差を右側に白色のグラフで示す。図 13 と同様、撮影位置によって色分けをしてある。また、背景が濃い灰色の部分は、4 ペア以上の対応点が見つからないため、方位推定に失敗した画像であることを示している。

7.3.3. 考察

図 13 と図 14 のグラフより、画像によって推定誤差にバラツキがあることがわかる。誤差の比較的大きい画像を観察すると、写っている地面領域が、

- 図 15 のように、車道など、単一色で特徴の少ない模様である
- 図 16 のように、バイクや看板など、地面領域のオクルージョンがある



図 15 : 特徴の少ない地面領域の例



図 16 : オクルージョンと経時変化のある地面領域の例



図 17 : 特徴が多く、推定のしやすい地面領域の例

- 図 16 のように、植物など、経時変化のある地面である

といった特徴が確認された。反対に精度の良かった画像には、図 17 のように、模様のある地面やマンホールなどが写り込んでいるといった特徴が見受けられた。これらから、SIFT による特徴量記述が容易な地面が写っているか否かが、推定精度に影響することが考えられる。

図 13 に示す位置推定の結果において、2 番や 17 番のように大きな誤差が発生した画像もあるが、黒線で示した一般的な GPS 測位の誤差 7.5[m]と比較すると、おおむね有効性が確認できる。

図 14 に示す方位推定の結果において、1 番や 7 番、24 番のように、SIFT を用いた対応探索がうまくいった場合には高精度の推定ができていたが、誤った対応探索結果により、60 度以上の誤差が出てしまったものや、方位推定に失敗してしまっただけのものも多く、方位推定に関しては電子コンパスと比較して有効性を見出すことが難しかった。

8. おわりに

本論文では、モバイルカメラで撮影した画像から、仮想的に視点を俯瞰視点へと変換した仮想俯瞰画像を生成する手法と、その仮想俯瞰画像と、高所から地面を鉛直に撮影した俯瞰画像との間で SIFT を用いた対応点探索を行うことにより、モバイルカメラの位置・方位を推定する手法について述べた。実験を通じ、提案手法の有効性を確認した。

今後の課題としては、ユーザ介入の必要な撮影条件の緩和やチルト角測定の自動化、方位推定の高精度化、解像度の低い空撮画像での実装などが挙げられる。

参考文献

- [1] 安田明生, “GPS 技術の展望”, 電子情報通信学会論文誌 B, Vol.J84-B No.12, pp.2082-2091, 2001.
- [2] 大友功一, 谷口哲司, 岸本正, 大山毅, “自律走行トラクタに関する研究-電子コンパスによる方位の計測-”, 帯広畜産大学研究報告 21, pp15-21, 1999.
- [3] “GoogleEarth,” Google, [オンライン]. Available: <http://www.google.co.jp/intl/ja/earth/index.html>. [アクセス日: 15 3 2012].
- [4] H. Kato and B. Mark, "Marker Tracking and HMD Calibration for a Video-based Augmented Reality

Conferencing System," In Proc. of the 2nd International Workshop on Augmented Reality (IWAR), 1999.

- [5] 水流弘達, 北原格, 大田友一, “環境に設置したステレオカメラを用いたモバイルカメラの位置・姿勢推定法,” 電子情報通信学会技術研究報告, vol. 111, no. 235, MVE2011-34, pp. 17-22, 2011.
- [6] 野田雅文, 高橋友和, 出口大輔, 井手一郎, 村瀬洋, 小島祥子, 内藤貴志, “空撮画像と時系列車載カメラ画像との照合による自車位置推定,” 電子情報通信学会技術報告, PRMU, パターン認識・メディア理解 Vol.109 (No.306), pp177-182, 2009.
- [7] D. G. Lowe, "Distinctive Image Features from Scale-Invariant Keypoints," International Journal of Computer Vision (IJCV), pp91-110, 2004.
- [8] R. Hertley and A. Zisserman, Multiple View Geometry in computer vision, 2nd Edition, Cambridge University Press, pp.32-36, 2003.
- [9] 藤吉弘亘, “一般物体認識のための局所特徴量(SIFT と HOG),” PCSJ/IMPS ナイトセッション, 2008.
- [10] M. A. Fisher and R. C. Bolles, "Random sample consensus:A paradigm for model fitting with applications to image analysis and automated cartography," Comm.of the ACM, Vol.24, pp.381-395, 1981.
- [11] “電子国土ポータル,” 電子国土事務局, [オンライン]. Available: <http://portal2.cyberjapan.jp/site/mapuse/index.html>. [アクセス日: 15 3 2012].