

実画像と 3D 地図間のカメラ位置推定に関する研究

入江 徹*, 藤 耕平*, 内海 公志*

*株式会社ジオ技術研究所 研究開発部

概要: 我々は、3次元化した市街地図を作成するにあたり、ビデオ(ハイビジョン)画像でのデータ収集を行っている。この収集された画像から建物の形状復元やテクスチャ切り出しを行うために、フレーム単位での正確な画像位置を知る必要がある。しかし、現在の GPS やベースとなる 2D 地図の精度により、収集した調査画像と 3D 地図上の位置を合わせるには多くの労力を必要としている。本論文では、調査画像と 3D 地図間の対応関係を入力とし、最適化問題に帰着させてカメラ位置を求める手法と、その処理を行うシステムについて述べる。

Camera Position Estimation between Real Image and 3D Map

Tohru Irie*, Kouhei Tou*, and Masashi Uchinoumi*

*Dept. of R&D of GEO Technical Laboratory Co., Ltd.

Hakata Gion Bldg, 1-1 Gion-machi Hakata-ku, Fukuoka, 812-0038 Japan

Abstract: Our products are 3D city map using real images (called research images). Here, it is important to know where the image was snapped, due to the 3D building making or texture extraction. However, it is difficult to match the viewpoint between research image and 3D map, because of the precision of GPS or 2D base map. This paper describes the optimization method of camera position using Simulated Annealing (SA).

1. はじめに

近年、カーナビ、防災・都市計画、アミューズメント等、様々な目的に対して GIS を用いたシステムが構築され実用化されている。当初、技術的・経済的な理由から 2 次元的に表現されていたこれらのシステムもハード・ソフトの性能向上に伴い、当然の流れとして 3 次元空間上で表現されるようになり、将来の介護・福祉・防災ロボット事業も含め、その適用分野は更に幅広く加速している。

我々は、(株)ゼンリンの 2 次元地図データを元に Walk eye Map と呼ばれる 3D 地図データの開発を行っている。建物形状のベースとなるのは、(株)ゼンリンの 2 次元地図データであり、建物等のテクスチャは、図 1 に示す計測専用車輜(タイガーアイ)を用いて正面・側面画像を HD カメラで撮像している。次に、撮影されたテクスチャを切り出して 3D の建物に貼り付けて街並の作成を行う。このとき、地表の隆起や、高速道路の高架の交差状況も考慮したモ

デリングが行われている。データ収集の概略図を図 2 に、モデリングされた画像例を図 3 に示す。

3D 地図を作成する際に、道路付帯物の位置・建造物高さの確認、自転車位置取得等のために、実世界で撮像された画像が 3D 空間のどの視点に対応するかを正確に知る必要がある。街並を撮影する計測専用車輜には、市販の GPS が積載されており、緯度・経度・高さが取得できるが、商用の衛星を用いた市街地等の高層建造物に囲まれた状態での移動体に対する位置測定の精度はメートルのオーダーである(注: GPS 測量等で用いられる干渉測位方式では固定 2 地点における GPS 受信機での複数回の観測によって数 mm から数 cm の精度で測量することが可能である^[1])。このため、厳密な意味で実世界での撮像した位置を GPS の軌跡から 3D 空間上で正確に反映することは困難であり、従来、オペレータが 3D 空間での視点を手動で変更しながら実画像との位置合わせを行っていた。本論文では、この視点の位置推定を最適化問題に帰着させ、シミュレーテッドアニーリング^[2](焼

きなまし法，以降SAと表記)を用いて最適化を行う手順と，それを実現するためのOpenGLを用いたシステムの概要について述べる。



図1 計測専用車輛(タイガーアイ)

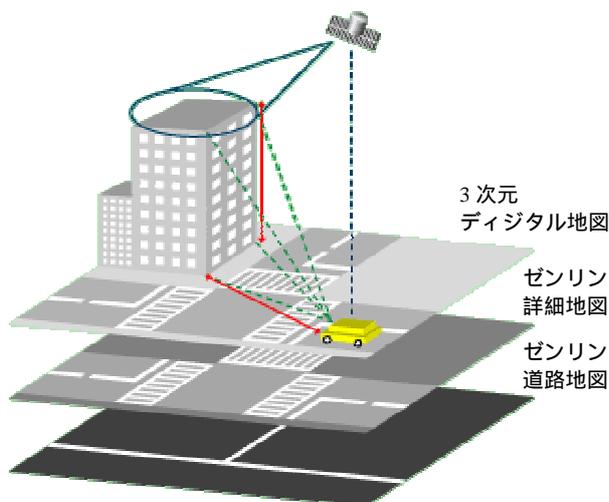


図2 データ収集概略図



図3 製品画像例

2. カメラ位置推定システム

2.1 システム構成

本システムは，各種センサーとハイビジョンカメラを搭載した計測専用車輛から得られたリサーチ画像と，(株)ゼンリンの2次元デジタル地図をデータソースとして立ち上げられた3次元建造物との間の位置合わせを行う(図4)。

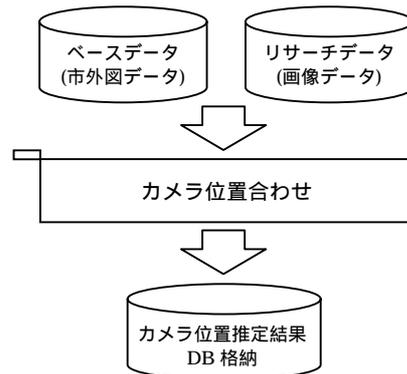


図4 データフロー

2.2 インターフェース概観

本システムのインターフェース概観を図5に示す。以降，それぞれのウィンドウを2D，3D，リサーチ画像面と呼ぶ。

2D 画像面(図6)は，地図(下地面)を真上からの視点で描画しており，同時に航空写真を透過画像で重ね合わせることができる。図6で示した2D 画像面において，中心の上三角形がカメラ位置であり，周辺の下三角形は，手動によって動画から切り出されたリサーチ画像の大まかな位置である。このリサーチ位置をクリックすることでリサーチ画像を選択することができる(一覧からの選択も可能)。

リサーチ画像面に表示されるのは，計測専用車輛のHDカメラによって撮像された画像である(図7)。

3D 画像面は2D 画像面とは別視点で描画され，建物は階数情報により高さ・種別毎に色を変えて簡易建物が表示される(図8)。また，リサーチ画像を透過画像で重ね合わせることができる。このリサーチ画像面と，3D 画像面で得られる視点が近くなるように3D 画像面のカメラ位置を調整することが目的となる。



図5 システム概観

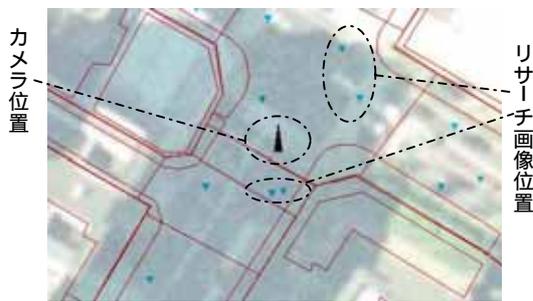


図6 2D 画像面拡大図



図7 リサーチ画像面

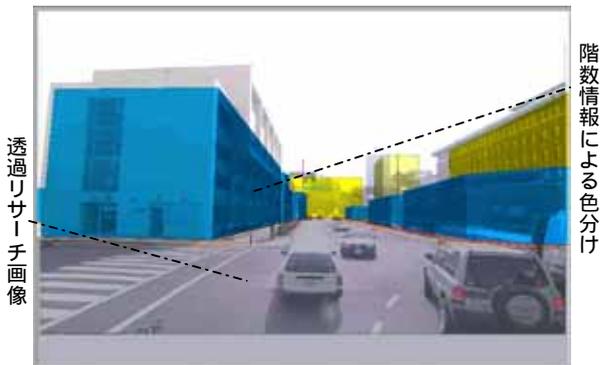


図8 3D 画像面

2.3 各ウィンドウ間の射影関係

図5に示した各ウィンドウ間の射影位置関係について図9に示す。図9において、2D画像面(*near*面)で指定された座標(U_m, V_m)と視点を結ぶ直線と、視体積限界面(*far*面)との交点が3次元座標(X, Y, Z)となり、その3次元座標が3D画像面に射影される座標が(U_c, V_c)となる。この3D画像面に映る画像がリサーチ画像と同一視点になるように視点2のカメラパラメータを変更する。つまり、3D画像面上での座標位置(U_c, V_c)と、リサーチ画像面上での座標位置(U_r, V_r)間の座標値の差を最小にするカメラパラメータ P を SA を用いて推定することになる。

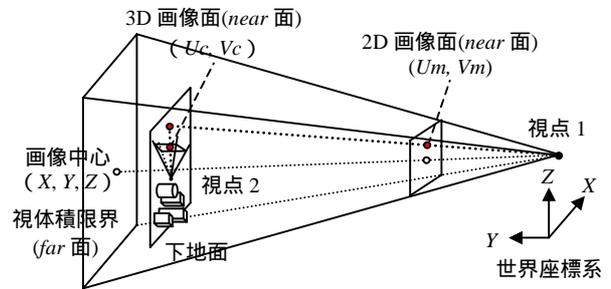


図9 2D-3D 画像座標位置関係

2.4 対応点入力方法

2.3で述べたように、リサーチ画像面と3D画像面間の対応点座標の差を最小にするようなカメラ位置を推定する。そのため、オペレータが両画像における対応点对を3対以上入力する必要がある。対応点の入力は、図5のリサーチ画像面 - 3D画像面間、またはリサーチ画像面 - 2D画像面間で行う。入力となる対応点として図10に示すような例が挙げられる。

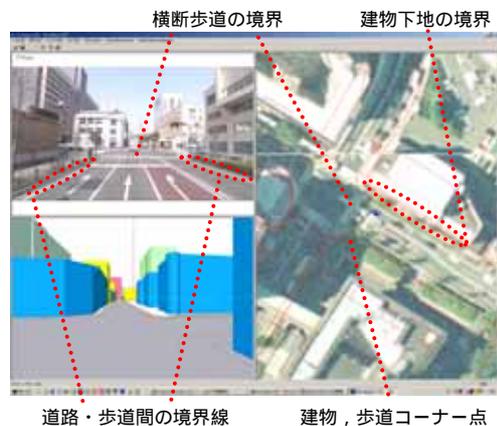


図10 入力点となる特徴

以上を踏まえて、対応点として入力する点を以下のように定義した(図 11)。

- 1) 特異点数は 3 点以上で、空間の広い範囲(奥から手前に向かう等)で道路下地・歩道下地・建物下地の境界線に沿う線分とする。
- 2) 1)が存在しない場合は、横断歩道境界線、建物・歩道コーナー点を入力とする。

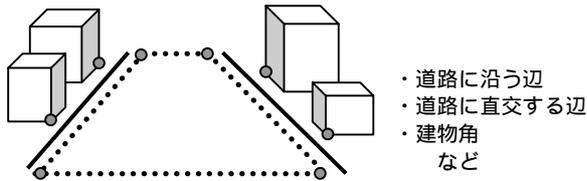


図 11 入力点候補となる点・線分

3. SA について

3.1 SA の概略

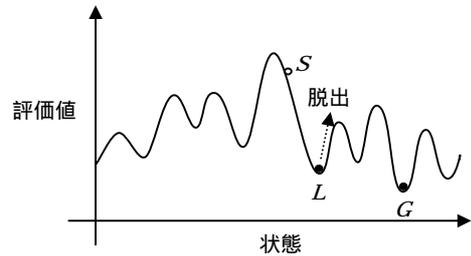
SA は、初期状態から出発し、より良い解を目指してその状態の局所近傍 $N(S)$ を探索する。現在の状態 (S) から出発して局所近傍の状態がそれより良い評価値を持つならばその状態へ遷移する。これを図 12 に示す。図 12 の S に示される初期解から出発し、 L に示す局所的最小に到達する。しかし、局所的最小 L から G に示す大域的最小に到達するには現在の解よりも悪い解を採択しなければならない SA では温度パラメータ (T) を設定し、改悪解の選択は温度が高いほど数多く行われる。繰り返し回数が増えるに従って温度パラメータは低下していき、それに伴って改悪解の選択回数は減少していく(絶対零度で SA は欲張り法と同一となる)。

解の選択確率は以下の式で決定される

$$P(S_j) = \begin{cases} 1 & \Delta Cost = (NCost_j - CCost_i) \leq 0 \\ e^{-\frac{\Delta Cost}{T}} & \Delta Cost = (NCost_j - CCost_i) > 0 \end{cases} \quad (1)$$

ここで、 $C\text{Cost}$ 、 $N\text{Cost}$ はそれぞれ現在の状態のコスト、新状態のコストであり、コスト最小化する場合の式である。(1)式において、新状態のコストが現在の状態のコストを改善する場合($N\text{Cost} - C\text{Cost} \leq 0$)、その採択確率は 1 であり、必ず採択される。解が改善されない場合($N\text{Cost} - C\text{Cost} > 0$)についても、確率 $e^{-\frac{\Delta Cost}{T}}$

で採択される。そしてこの(悪解)採択確率は、温度パラメータ T を伴って、温度減少と共に採択確率も減少していくことになる。



L : 局所的最適解
 G : 大域的最適解
 山登りが必要 $\left\{ \begin{array}{l} T \text{ 高: 山登り回数多} \\ T \text{ 低: 山登り回数少} \end{array} \right.$

図 12 SA における局所的最適と大域的最適

3.2 カメラ位置推定への適用

前述の SA を用いて推定するのは 3D 空間でのカメラ中心を表す $Eye(X, Y, Z)$ 座標、画像中心を表す $Center(Cx, Cy, Cz)$ 、Up vector を表す (Ux, Uy, Uz) である。その位置関係を図 13 に示す。

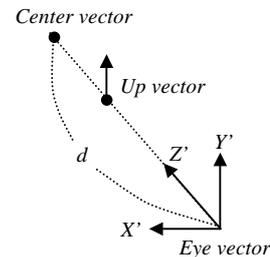


図 13 カメラ座標系

各パラメータの探索範囲は、実験により以下のように設定する。

$$Eye \text{ vector} : -200 \leq \text{現在値} \leq 200. \quad (2)$$

$$Center \text{ vector} : -200 \leq \text{現在値} \leq 200. \quad (3)$$

$$Up \text{ vector} : -0.01 \leq \text{現在値} \leq 0.01. \quad (4)$$

ここで、Up vector の値域は $-1.0 \sim 1.0$ であるため、(4)のように設定した。実際には(5)、(6)式の制約が存在するため、推定パラメータは 7 つとなる。

• Eye-Center 間の距離の制約: Eye-Center 間の距離は一定(距離を d とする)。

$$(Cx-X)^2 + (Cy-Y)^2 + (Cz-Z)^2 = d^2. \quad (5)$$

• Up vector と Eye-Center 間の直交性: Up vector と Eye-Center は直交。

$$Ux \cdot (Cx-X) + Uy \cdot (Cy-Y) + Uz \cdot (Cz-Z) = 0. \quad (6)$$

SA の初期温度は 1000 度，温度減少係数 α は 0.95 として(7)式により冷却を行う．

$$T = \alpha T. \quad (7)$$

パラメータの評価値は(7)式を用いて，図 9 におけるリサーチ画像面座標(U_r, V_r)と 3D 画像面座標(U_c, V_c)との距離の和が最小になるように求める．

$$\min \sum_{i=0}^N \{(U_{r_i} - U_{c_i})^2 + (V_{r_i} - V_{c_i})^2\}. \quad (8)$$

ここで， N は両画像間で対応する座標値の数である．

本手法では，より良い解を求めるため，前述の SA(10000 回)で得た解を再び SA の初期値として計算を繰り返す(5 回)．その際，(9)式を用いて，探索する近傍範囲を SA 適応回数に応じて縮小していく．

$$\text{探索範囲} = \text{探索範囲} / (\text{SA 適応回数}). \quad (9)$$

4. 評価実験

4.1 実験環境

検証実験として，実際のリサーチ画像を用いてカメラの位置推定を行った．開発・実験環境は以下の通りである．

Computer: DELL 社製 Precision 340,

CPU: Intel 社製 Pentium4, 1.60GHz,

Memory: 1024MB,

開発環境: Microsoft Visual C++ 6.0.

4.2 実験結果

4.2.1 解の収束結果

SA による最適解探索の過程を示すために，カメラ位置推定を行った際の(8)式の評価値の変化を見る．図 14 に，実験に用いたリサーチ画像を示す．ここで，対応点として入力した点を図中の で示している．

図 14 の画像に対し，初期値を変更して最適解探索を 30 回試行した際の評価値の変化を図 15 に示す．ここで，評価値 6000，計算回数 2000 を越えるものについては表示を省略している．また，図 16 に，各試行における初期評価値と最終評価値を示す．加えて，図 17 に各試行の計算時間を示す．

図 15, 16 より，一部の場を除外して初期値の変化

に関わらず評価値を最小(と思われる)にする解に収束していることが分かる(一部は局所最適と思われる解に収束している)．また，計算時間について 1 秒以下で安定的に処理を終えていることが分かる．

図 18 にカメラ位置推定を行い，3D 画像面とリサーチ画像面を重ね合わせた結果例を示す．

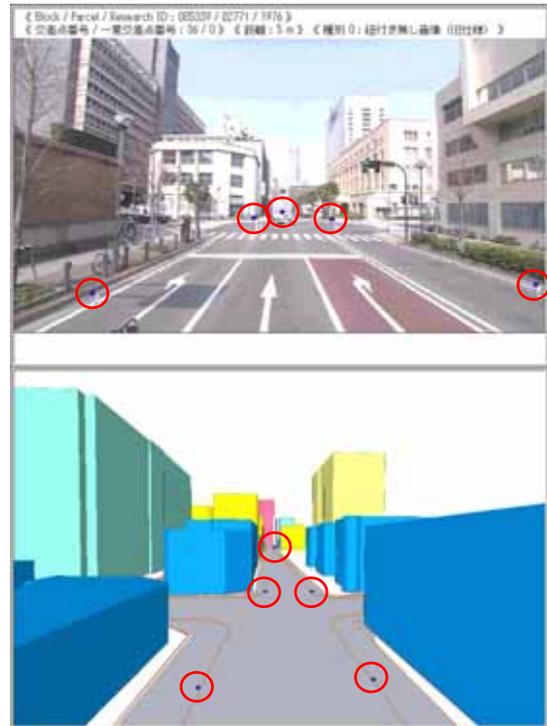


図 14 リサーチ画像と入力点

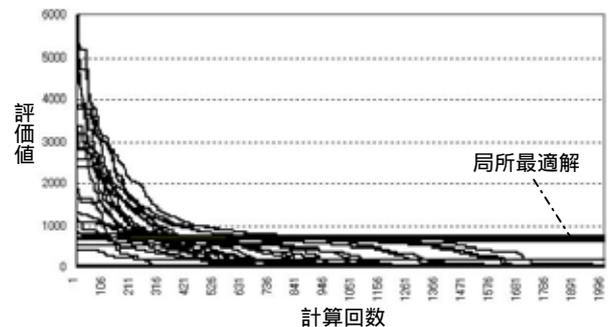


図 15 評価値の変化(評価値 6000 以上省略)

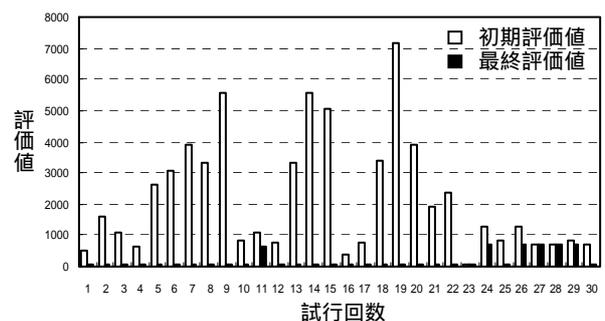


図 16 初期評価値と最終評価値

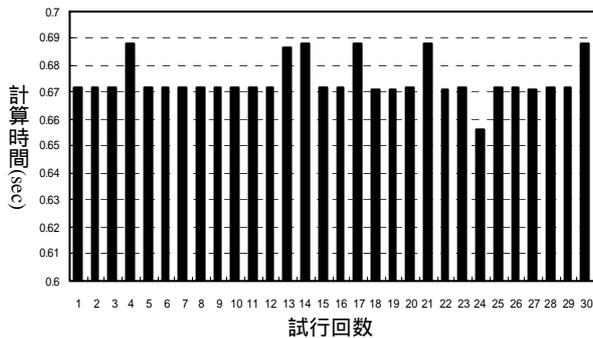


図 17 各試行の計算時間

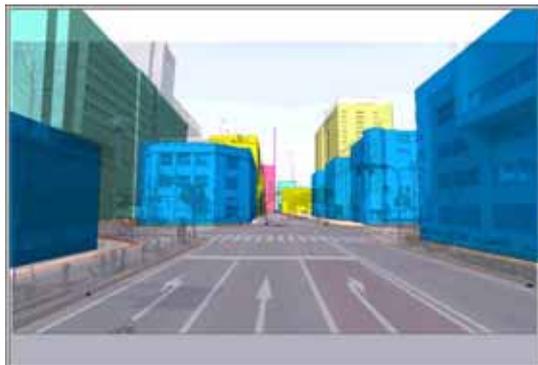


図 18 カメラ位置推定結果例

4.2.2 カメラ位置推定結果

次に、複数の異なるリサーチ画像(サンプル: 1 交差点分 21 枚)に対して提案手法を適用した結果について述べる。図 19 に、使用したリサーチ画像の一部を示し、図 20 に各画像に対する初期・最終評価値を示す。また、図 21 に各画像のカメラ位置合わせの計算時間を示す。

図 20 より、一部の画像を除いて(8)式で表わされる入力座標値の差を最小にするように収束していることが分かる。一部の評価値が大きい画像(画像番号 3, 6, 10, 13, 14, 16 等)については、計算終了時のパラメータを初期値として再び最適化処理を行うことで、評価値を下げる事ができた。これは繰返し計算回数を増やすことで解消できるが、実用的には一度の試行後、位置合わせの確認を行い、不十分であれば再度試行するという手順を採用している。

5. まとめと今後の課題

実画像と 3D 地図間のカメラ位置推定に最適化アルゴリズムの 1 つである SA を用い、その有効性を検証した。位置合わせ精度、計算時間面でバランス

のとれた最適化を実現することができた。今後の課題として、現在オペレータによる対応点入力の自動化が挙げられる。そのためには 3D モデルと実画像から幾何形状などの特徴を抽出し、マッチングを行う必要がある。

参考文献

- [1] 坂井丈泰, GPS 技術入門, 東京電機大学出版(2003)
- [2] Sadiq M. Sait and Habib Youssef, 組合せ最適化アルゴリズムの最新手法 基礎から工学応用まで, 丸善(2002)
- [3] 廣瀬通孝, 渡辺真二郎, 谷川智洋, 遠藤隆明実写画像を用いた広域仮想空間構築における画像生成手法の研究, 信学技報, MVE98-32, pp.53-58 (1998)



図 19 交差点周辺リサーチ画像サンプル

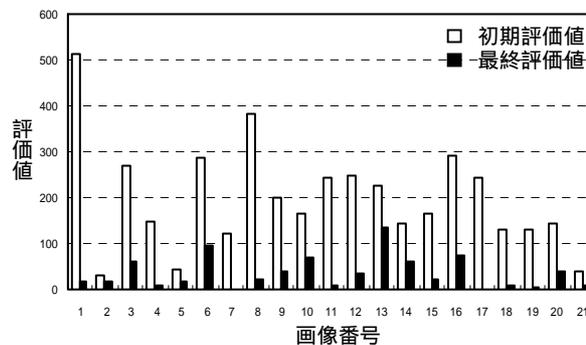


図 20 初期評価値と最終評価値(1 交差点)

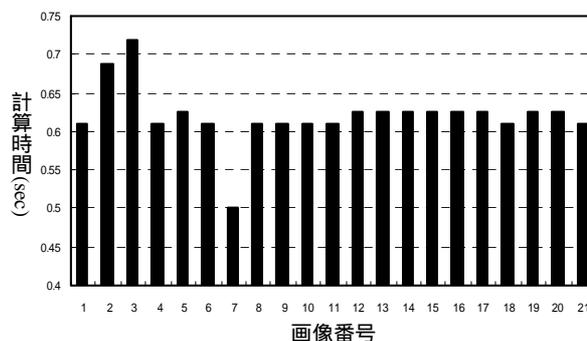


図 21 各試行の計算時間(1 交差点)