

# カメラ間の位置関係に基づく画像間距離系列を用いた 車載カメラ映像データベース検索による自車位置推定

非会員 久徳 遙矢<sup>\*a)</sup> 非会員 出口 大輔<sup>\*\*</sup> 非会員 高橋 友和<sup>\*\*\*</sup>  
非会員 目加田慶人<sup>\*\*\*\*</sup> 非会員 井手 一郎<sup>\*</sup> 非会員 村瀬 洋<sup>\*</sup>

## Ego-localization by Searching in a Streetview Database using a Sequence of Image Distance based on the Positional Relation between In-vehicle Cameras

Haruya Kyutoku<sup>\*a)</sup>, Non-member, Daisuke Deguchi<sup>\*\*</sup>, Non-member, Tomokazu Takahashi<sup>\*\*\*</sup>, Non-member,  
Yoshito Mekada<sup>\*\*\*\*</sup>, Non-member, Ichiro Ide<sup>\*</sup>, Non-member, Hiroshi Murase<sup>\*</sup>, Non-member

(2013年9月25日受付, 2014年2月7日再受付)

We propose an accurate ego-localization method by searching a streetview database composed of single front-view in-vehicle camera images. Previously, we proposed an image distance metric reflecting the positional relation of two cameras based on epipolar geometry analysis, and used it for ego-localization. However, since the method employed a dynamic time warping strategy to avoid the effect from outliers, both input and database images needed to be image sequences. To overcome this problem, the method proposed in this paper reformulates the previous image distance metric to a novel image distance that requires only a single in-vehicle camera image as an input, which is realized by considering the sequential property of the images in the database. We conducted experiments using multiple image sequences captured under various conditions by using an in-vehicle camera mounted on the windshield of a car. The experimental results showed that the proposed method could achieve an accuracy of 89%, and we confirmed its effectiveness.

キーワード：車載カメラ, 自車位置推定, ITS

**Keywords:** In-vehicle camera, Ego-Localization, ITS

### 1. はじめに

本論文では、自動車における運転者支援システムに着目する。我々にとって身近な運転者支援システムとして、カーナビゲーションシステムがある。これは、自車位置情報と地図情報を元に、運転者を目的地まで案内するシステムである。カーナビゲーションシステムの2012年末時点での出荷累計台数は5,400万台を超えており<sup>①</sup>、快適な運転を支援するために欠かせないシステムである。近年では、スマートフォンなどを用いたポータブルナビゲーションデバイスの普及も進んでおり<sup>②</sup>、ナビゲーションシステムがより身近な存在となりつつある。さらに、これらのデバイスから得られる自車位置情報は、ナビゲーションのためだけでなく、震災における迅速な交通状況の把握<sup>③</sup>、事故が発生しやすい地点の抽出<sup>④</sup>、といった様々な用途で用いられている。このような背景から、自車位置情報を得るための技術に対する需要は高い。

自車位置を得るためのセンサとして、普及型GPSが広く一般に用いられている。しかし、普及型GPSにより得ら

a) Correspondence to: Haruya Kyutoku. E-mail: kyutoku@murase.m.is.nagoya-u.ac.jp

\* 名古屋大学 大学院情報科学研究科  
〒464-8601 愛知県名古屋市千種区不老町  
Graduate School of Information Science, Nagoya University  
Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

\*\* 名古屋大学 情報連携統括本部  
〒464-8601 愛知県名古屋市千種区不老町  
Information and Communications Headquarters,  
Nagoya University

Furo-cho, Chikusa-ku, Nagoya, Aichi 464-8601, Japan

\*\*\* 岐阜聖徳学園大学 経済情報学部  
〒500-8288 岐阜県岐阜市中鶯 1-38

Faculty of Economics and Information,  
Gifu Shotoku Gakuen University

1-38, Nakauzura, Gifu, Gifu 500-8288, Japan

\*\*\*\* 中京大学 工学部

〒470-0393 愛知県豊田市貝津町床立 101  
School of Engineering, Chukyo University  
101, Tokodachi, Kaizu-cho, Toyota, Aichi 470-0393, Japan

れる位置情報には、市街地で数 m、高層ビル街では数 10 m の誤差が含まれる<sup>(9)</sup>。それゆえ、より高度な運転者支援のために、高精度に自転車位置を取得する技術が求められている<sup>(9)</sup>。

一方、Google ストリートビュー<sup>†</sup>に代表されるように、位置情報付きの車載カメラ映像からなる市街地映像データベースも整備されつつある。近年、このような市街地映像データベースを利用した自転車位置推定手法が注目を集めている。これは、車載カメラから得られた画像に対応するデータベース中の画像を検索し、その画像に付随する位置情報を参照することで自転車位置を求めるものである。これらは、位置推定に用いる入力データとして、映像(画像系列)を用いるものと、単一の画像のみを用いるものの2つに分けられる。

位置推定に映像を用いる手法として、入力に全方位カメラ映像を用いる手法<sup>(10)</sup>、斜め側方を向いた通常カメラ映像を用いる手法<sup>(11)</sup>が提案されている。これらは照合に領域ベースの画像特徴を用いるため、他車両や街路樹などによる遮蔽に弱い。この問題を解決するため、我々は汎用性が高い正面向きカメラを入力とする特徴点ベースの自転車位置推定手法を提案した<sup>(9)</sup>。この手法は、入力およびデータベース共に正面向き車載カメラを用い、遮蔽や走行位置の変化に伴う見えの変化に頑健な画像間距離を用いることで、高精度な自転車位置推定を実現した。しかし、これらの手法は車載カメラ映像とデータベース映像の DP マッチングにより自転車位置を求めるため、DP マッチングの端点を精度良く与える必要がある。また、高精度に端点を与えない場合、DP マッチングの計算コストが著しく増加するという問題点もある。

一方、カメラにより撮影された1枚の画像のみを用いて市街地映像データベースと照合し、位置推定を行なう手法も提案されている。山崎らは、エッジ特徴を用いた位置推定手法を提案している<sup>(10)</sup>。しかし、障害物の回避行動に伴う走行位置の変化や、大きな遮蔽による見えの変化を考慮していない。薄らは、ランドマークデータベースと入力画像を SIFT<sup>(11)</sup>により対応付ける手法を提案している<sup>(12)</sup>。しかし、ランドマークデータベースの構築には、全方位カメラ映像中の各特徴点の3次元位置を求める必要があり、大規模なデータベースへの適用は困難である。また上野らは、SIFT の対応点数を尺度として市街地映像データベースとの照合を行なう手法を提案している<sup>(13)</sup>。この手法は、画像中の建物領域の平面性に注目し、幾何的な拘束を用いて精度改善を行なっている。しかし、建物のように大きな平面を持つ物体が多く含まれる地点以外への適用は困難である。

以上の背景から、本論文では車載カメラで取得した単一の画像を入力とし、市街地映像データベースとの高精度な対応付けにより自転車位置を推定する手法を提案する。ここで、入力およびデータベース共に正面向き車載カメラを用い、どちらも道路方向と車両進行方向が平行な状態で撮影されていると仮定する。車載カメラは事故などのトラブル

の記録を目的として設置されることが多いため、正面向きの車載カメラが最も実用的かつ汎用的であると考えられる。一方、正面向き車載カメラから得られた画像間の対応付けは、遮蔽や走行位置の変化に伴う見えの変化の影響を受けやすい。さらに、撮影条件が異なる場合は、従来の画像間距離では安定して画像の対応を求めることが困難となる。そこで提案手法では、データベースが道路に沿って撮影された画像系列である点に着目し、以前我々が提案したカメラの幾何的な位置関係を考慮した見えの変化に頑健な画像間距離<sup>(9)</sup>を拡張する。これにより、画像の撮影条件が異なる場合に画像間距離が安定して求まらないという問題を解決する。また、文献(9)において自転車位置推定の入力に画像系列が必要であるという制約を緩和し、1枚の車載カメラ画像のみを入力して自転車位置推定可能な手法を実現する。

以降、2章で我々が以前提案した画像間距離<sup>(9)</sup>および本論文で提案する新たな画像間距離について説明し、3章で提案手法の具体的な計算手順について述べる。そして、4章で実験について説明し、5章にて提案手法の有効性について議論する。最後に6章で本論文をまとめる。

## 2. データベースの時系列情報を利用した画像間距離

本章では、まず(2.1)で、我々が文献(9)で提案した、正面向きカメラにより撮影された2枚の画像間の距離について説明する。これは、エビポーラ幾何の概念<sup>(14)</sup>に基づいており、遮蔽や走行位置の変化に伴う見えの変化に頑健な尺度となっている。そして、本論文で新たに提案する画像間距離について(2.2)で詳細に説明する。

### (2.1) 2カメラの幾何配置を考慮した1対1画像間距離

入力である走行中の正面向き車載カメラ画像を  $A$ 、道路に沿って撮影された正面向き車載カメラ映像データベースを  $B = \{B_i\}$  とする。ここで  $B_i$  は、データベースである画像系列における  $i$  番目の画像である。本研究では、入力画像  $A$  と最も近い位置で撮影されたデータベース中の画像  $B'$  を検索して求めることで自転車位置推定を行なう。しかし、同じ道路で撮影された車載カメラ画像同士であっても、正面向き車載カメラから得られる画像は走行軌跡の違いによる見えの変化が大きく、拡大・縮小・回転のみを考慮した単純な画像照合では精度よく画像間の対応を得ることができない。また、前方車両や路上駐車車両などによる遮蔽の影響も強く受ける。

上述の問題を解決するため、文献(9)ではカメラ間の幾何的な位置関係を反映した画像間距離を提案した。具体的には、入力およびデータベース共に道路方向と車両進行方向が平行な状態で撮影されている、すなわち2カメラの光軸は平行であると仮定する。この仮定の下、Fig. 1 のような光軸が平行である2つのカメラ位置とエビポール位置の関係から画像間の距離を定義した。エビポールは一方のカメラから見た他方のカメラの画像平面上での位置を表すことから、Fig. 1 (a) のように光軸方向に2カメラが離れている場合はエビポールと画像中心の距離は小さく、Fig. 1 (b)

<sup>†</sup> Google ストリートビュー, <http://maps.google.co.jp/>

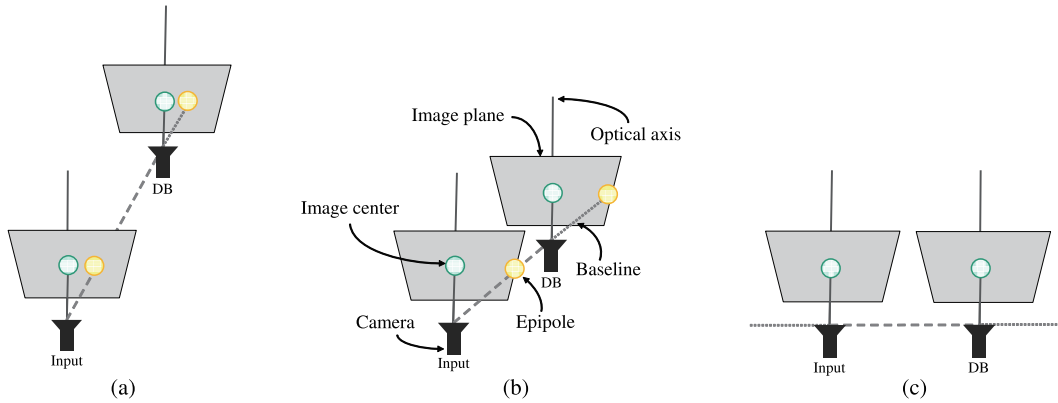


Fig. 1. Relation between the positions of two cameras and epipoles.

のようにカメラが近付くとその距離は大きくなる。さらに、Fig. 1 (c) のような位置関係の場合、つまり 2 カメラ間の距離を進行方向に沿って測った値が小さいとき、エピポールと画像中心の距離は無限大に近くなる。

Fig. 2 に、車載カメラから得られた入力画像  $A$  と、データベース内の最も近い位置で撮影された画像  $B'$ 、およびその前後の画像に対し、実際にエピポールの位置を算出して描画したものを示す。図中の直線はエピポーラ線を表わしており、それらの交点がエピポールである。図から、2 カメラ間の距離、つまり  $A$  に対応する  $B'$  からの車両進行方向のずれに応じて、Fig. 1 と同様にエピポールの位置 (図中の丸印) が変化することを確認できる。

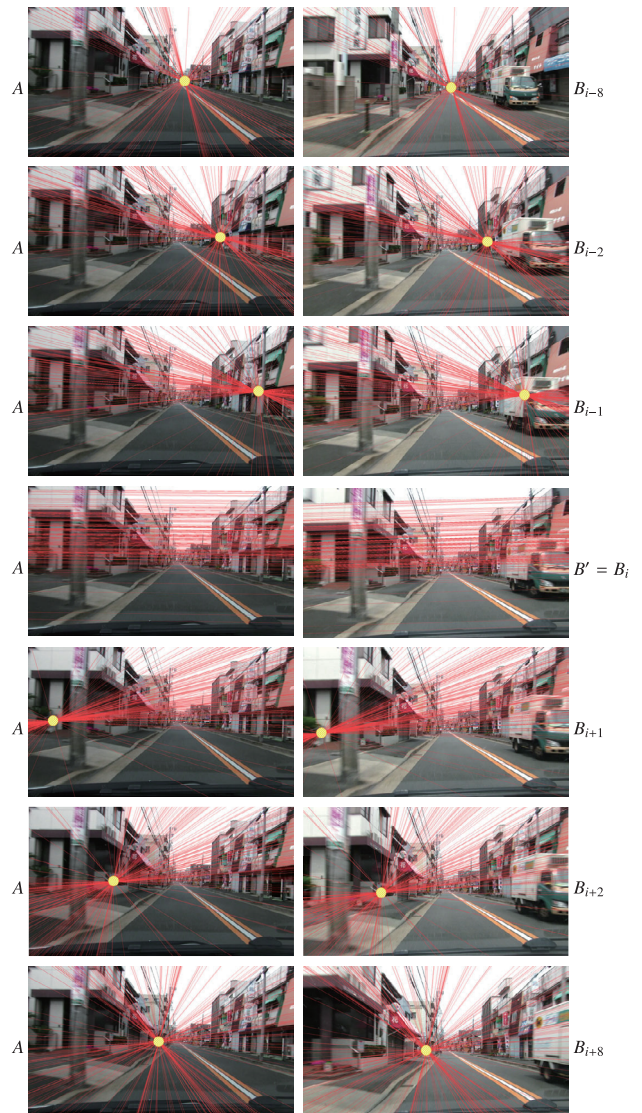
以上の性質から、文献 (9) では、エピポールと画像中心の距離の逆数を画像間距離とした。この画像間距離は 1 枚の画像同士から算出されるものであることから、以降では 1 対 1 画像間距離と呼ぶ。もし、常に正確なエピポールの位置を算出できるならば、入力画像に対応するデータベース中の画像は、この 1 対 1 画像間距離が最小となる画像である。しかし、実環境における車載カメラから取得した画像から常に正確なエピポールの位置を求めることは困難であり、対応する画像とは離れた位置で画像間距離が最小となる可能性がある。そこで提案手法では、以下のようにデータベースの時系列情報を用いることで、外れ値に頑健な検索を実現する。

**〈2・2〉 1 対 1 画像間距離の対称性に基づく 1 対多画像間距離**

〈2・1〉で述べた画像間距離の性質を、より詳細に見る。入力となる車載カメラの位置を  $O$ 、データベース中のある画像を取得したカメラ位置を  $Q$ 、焦点距離を  $f$  とし、エピポールの位置を  $E$  とすると、各位置は Fig. 3 のような関係となる。 $x$  方向は車両の横方向に対する走行位置、 $z$  方向は車両の進行方向に対する走行位置を表わす。車両は道路平面上を移動するため、ここでは  $xz$  平面での位置関係のみに着目する。このとき、画像中心とエピポールの距離  $|e_x|$  と、2 カメラ間の進行方向に対する距離  $|q_z|$  の関係は、

$$\frac{|e_x|}{f} = \frac{|q_x|}{|q_z|} \dots \dots \dots (1)$$

となる。両辺に  $f$  を乗じると、



Line: Epipolar-Line, Circle: Epipole.

Fig. 2. The position of epipoles between the input frame  $A$  (left column) and its closest frame  $B' = B_i$  and adjacent frames in the database (right column).

$$|e_x| = f|q_x| \frac{1}{|q_z|} \dots \dots \dots (2)$$

が得られる。ここで、光軸が平行であることから、 $|q_x|$  は

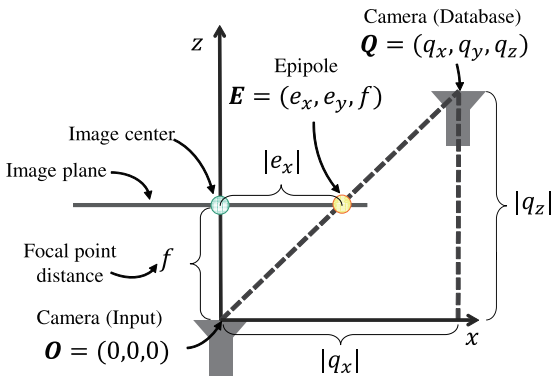


Fig. 3. Relation between the distance of the image center to the epipoles, and its relation with the gap between the cameras from the vehicle's direction.

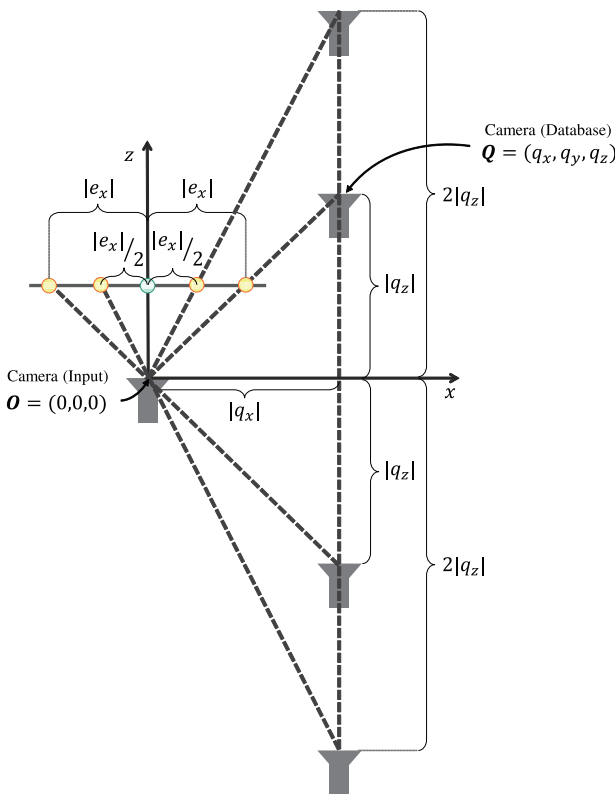


Fig. 4. Example of the epipole's position when the distances between the cameras are  $|q_z|$  and  $2|q_z|$ .

短時間で変化せず、焦点距離  $f$  が一定であるとする、 $|e_x|$  と  $|q_z|$  は反比例の関係となる。

以上から、1対1画像間距離、つまり画像中心とエピポールの距離の逆数  $1/|e_x|$  と、2カメラ間の進行方向に対する距離  $|q_z|$  の関係は、

$$\frac{1}{|e_x|} = \frac{1}{f|q_x|}|q_z| \dots\dots\dots (3)$$

となり、比例関係にあることがわかる。

一方、Fig. 4 はカメラの配置が Fig. 3 と異なる場合の例である。Fig. 4 から、入力とデータベースのカメラ配置が前後したとしても、進行方向に沿ったカメラ間の距離が  $|q_z|$  の場合は画像中心とエピポールの距離が  $|e_x|$  であることが

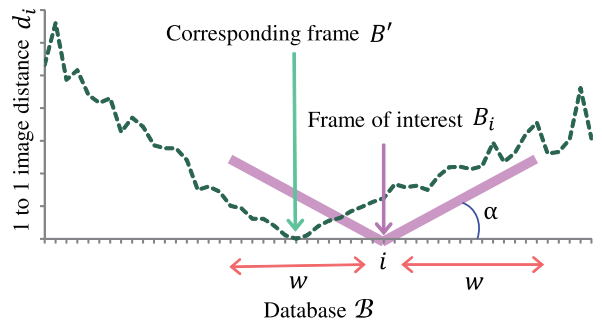


Fig. 5. Calculating the distance of images from a sequence of frame distances.

わかる。さらに、カメラ配置が前後のどちらであったとしても、進行方向に沿ったカメラ間の距離が  $2|q_z|$  の場合、画像中心とエピポールの距離は  $|e_x|/2$  となる。つまり、1対1画像間距離は、次の性質を持つことがわかる。

1対1画像間距離の性質  
1対1画像間距離  $1/|e_x|$  は、 $1/|e_x| \approx 0$  となるカメラ位置の画像を中心として対称な直線となる。

本論文では、走行中に得られる1枚の車載カメラ画像と、事前に構築した市街地映像データベースを用いて自車位置推定を行なう。ここで、市街地映像データベースがある時間間隔で撮影された画像系列であることから、入力画像1枚とデータベースの画像系列間で「1対1画像間距離の性質」を利用した距離を定義することができる。具体的には、データベースの画像系列における、進行方向のカメラ位置が等間隔で記録されているとみなせる区間、すなわちカメラが等速直線運動していると仮定できる短時間の区間に注目する。このとき、入力画像  $A$  とデータベース画像系列  $B = \{B_i\}$  の1対1画像間距離  $d_i$  は、入力画像  $A$  との進行方向に沿って測った距離が0となるデータベース画像  $B'$  を中心として時間方向に対称な直線となる。そこで、データベースのある画像  $B_i$  と、その前後  $w$  枚ずつの画像を含めた  $[B_{i-w}, \dots, B_i, \dots, B_{i+w}]$  を用い、 $A$  と  $B_i$  の1対多画像間距離  $D_i$  を次のように定義する。

$$D_i = \sum_{j=i-w}^{i+w} |d_j - \alpha|j - i|| \dots\dots\dots (4)$$

ここで、 $\alpha$  はデータベース中の注目画像  $B_i$  を中心とする対称な直線の傾きであり、右辺を最小にする値とする。Fig. 5 にこの概念図を示す。式 (4) の右辺が最小となる  $\alpha$  が与えられる時、式 (4) は入力画像  $A$  とデータベース画像  $B_i$  が「1対1画像間距離の性質」をどの程度満たしているかを表わす指標となる。そして  $B_i = B'$  のとき、すなわち  $A$  に対応するデータベース中の画像  $B'$  を中心としたとき、式 (4) は最小な値をとる。

以上のようにデータベースの時系列情報を利用した1対多画像間距離  $D_i$  を尺度として用いることで、個々の画像間距離だけでは判定できない外れ値の影響を抑えた検索が可能となる。

### 3. 1 対多画像間距離を用いた画像検索による自車位置推定

本章では、単一画像を入力とした高精度な市街地映像データベース検索による自車位置推定手法の具体的な計算手順について述べる。まず、位置情報付き市街地映像データベースをあらかじめ構築しておく。次に、1枚の車載カメラ画像が入力されたとき、以下の手順でデータベース内の対応する画像を検索する。

• **Step 1 : 対応点の算出**

入力およびデータベースの画像から、SIFT<sup>(11)</sup>を用いてキーポイントおよびその特徴量を求める。そして、それらの特徴量からキーポイント間の対応を求める。最後に、対応点群から RANSAC アルゴリズム<sup>(15)</sup>を用いて信頼性の高い対応点群を求める。

• **Step 2 : 1 対 1 画像間距離の算出**

Step 1 で求めた対応点群から基礎行列を求め、エピポールの位置を求める。そして、推定されたエピポールの位置を利用して、式 (3) の 1 対 1 画像間距離を算出する。

• **Step 3 : 1 対 1 画像間距離系列の算出**

入力画像とデータベース中の各画像において、式 (3) の 1 対 1 画像間距離を計算し、その系列を 1 対 1 画像間距離系列とする。

• **Step 4 : 1 対多画像間距離の算出**

1 対 1 画像間距離系列から、式 (4) により各画像位置を中心とした 1 対多画像間距離を求める。

• **Step 5 : 対応画像の決定**

1 対多画像間距離が最小となる位置の画像を、入力画像に対応する画像として選択する。

最後に、検索されたデータベース内の画像に付随する位置情報を出力する。

### 4. 実験

本章では、3章で述べた提案手法の Step 1 から 5 の、対

応画像検索精度に関する評価実験について述べる。まず実験に用いたデータを説明し、次に具体的な実験方法として、正解位置とする対応画像の決定方法、比較手法について述べる。最後に、実験結果について述べる。

〈4・1〉 **実験データ** 実験では、車載カメラとして CANON iVUS HF G10 を用い、このカメラを車両のフロントガラスに固定し、全長約 480m の道路を、可能な限り等速となるように 5 回走行して撮影した。撮影した映像は、1,920×1,080 画素、23.98 フレーム/秒である。実験に用いたデータの撮影条件 (天候、撮影時刻)、およびフレーム数を Table 1 に示す。また、取得した各系列の画像例を Fig. 6 に示す。

〈4・2〉 **実験方法** 実験用データ 5 系列のうち、ある 1 つの系列内の各画像を入力、残りの 4 系列中の 1 系列をデータベースとして用い、全ての系列の組み合わせに対して実験を行なった。具体的には、各入力画像をクエリとし、データベースからそれぞれの対応画像を求めた。比較手法として、文献 (13) で利用している、特徴点の対応点数を尺度とする手法を比較手法 1、2 章で述べた式 (3) の 1 対 1 画像間距離が最小となる位置を求める手法を比較手法 2 として用いた。なお、提案手法、比較手法 1、比較手法 2 のいずれの手法においても、3 章の Step 1 で求めた対応点群を用いた。また、提案手法の式 (4) では  $w = 3$  を用いた。

本実験では、正しい画像位置を選択できた割合を正解率として各手法の評価を行なった。ここで、対応付け誤差が 1 フレーム以内のものを正解とした。これは、路面上における ±35 cm 程度のずれに相当する。正しい画像間の対応

Table 1. Conditions of the experimental data.

Sequence	Condition	Number of frames
1	Light rain, daytime	1,413
2	Light rain, daytime	1,205
3	Fine, early morning	1,498
4	Cloudy, daytime	1,348
5	Fine, daytime	1,173

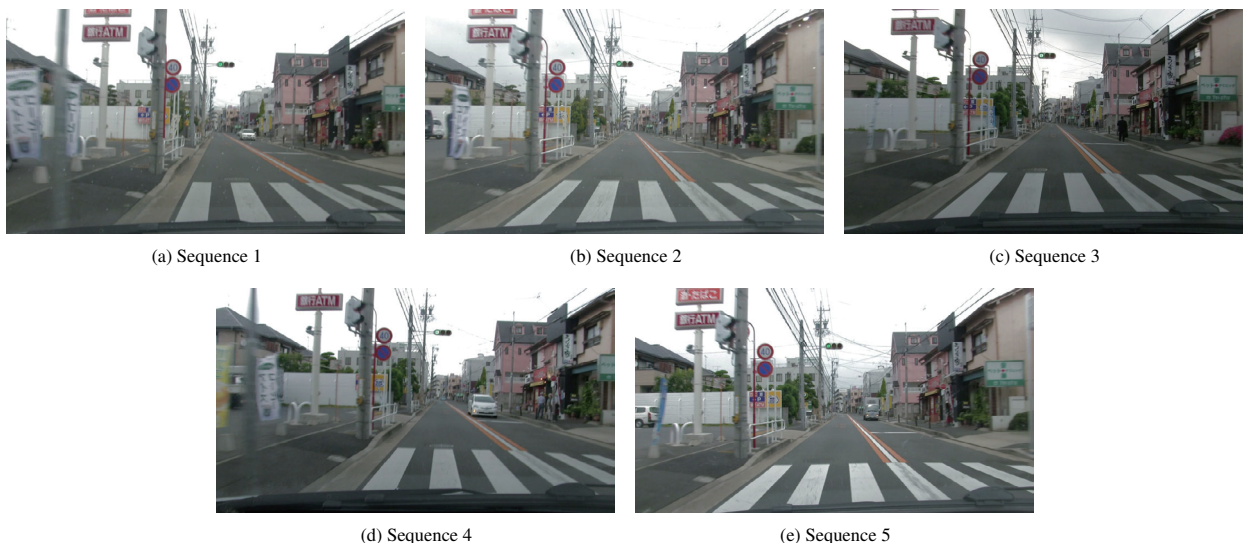


Fig. 6. Example of the experimental data.



Fig. 7. Search result of Sequence 1 from Sequence 2. The value in the bottom right of each image shows the difference (number of frames) from the correct corresponding image.

は、1対1画像間距離  $1/|e_x|$  と DP マッチングを用いる手法<sup>9)</sup>で求めた。この手法は誤差が平均 0.5 フレームと小さく、正解率も 99%と高精度であるため、正しい対応を与えるのに十分な性能を有していると考えた。

〈4・3〉 実験結果 各手法の正解率を Table 2 に示す。また、入力とデータベースの撮影条件が類似する場合の例として、系列 1 を入力、系列 2 をデータベースとしたときの検索結果例を Fig. 7 に示す。この組み合わせは、入力およびデータベース共に日中の小雨時に撮影された系列である。撮影時刻の違いにより画像の見えが大きく異なる例として、系列 5 を入力、系列 3 をデータベースとしたときの検索結果例を Fig. 8 に示す。いずれも、入力画像群から等間隔にサンプリングしたものに対する結果であり、左から入力画像、正解画像、そして提案手法、比較手法 1、比較手法 2 による検索結果である。各手法による検索結果の右下の数値は、推定誤差を正解画像からのフレーム数で表わしたものである。

## 5. 考 察

〈5・1〉 正解率の比較 Table 2 の全ての実験データに対する結果から、提案手法は比較手法 1 に比べて 42%、比較手法 2 に比べて 8%高い正解率が得られたことがわかり、提案手法の有効性を確認した。またいずれの系列の組み合わせにおいても、提案手法は比較手法 2 よりも高い正解率を得た。これは、提案手法により基礎行列推定失敗による外れ値の影響を避けて対応画像を求めることができたため

Table 2. Accuracy rates of corresponding image search.

Input	DB	Proposed	Comparative 1	Comparative 2
1	2	96%	28%	93%
	3	83%	25%	73%
	4	87%	40%	66%
	5	97%	50%	95%
2	1	94%	21%	91%
	3	85%	73%	80%
	4	88%	33%	75%
	5	95%	53%	83%
3	1	82%	28%	73%
	2	89%	85%	83%
	4	98%	56%	95%
	5	84%	56%	64%
4	1	84%	36%	67%
	2	97%	37%	93%
	3	99%	53%	97%
	5	86%	66%	83%
5	1	97%	43%	94%
	2	98%	55%	93%
	3	71%	41%	43%
	4	77%	60%	74%
All		89%	47%	81%

だと考えられる。本実験で正解と判断した 1 フレーム以内の誤差は、路面上における  $\pm 35$  cm 程度のずれに相当するため、普及型 GPS と比べて十分高精度に位置推定が可能であると言える。

〈5・2〉 撮影条件の違いに対する頑健性 小雨下で取得した系列 1、系列 2 をそれぞれ入力とデータベースとした場合、提案手法は比較手法 1 に比べ 68%高い正解率が得



Fig. 8. Search result of Sequence 5 from Sequence 3. The value in the bottom right of each image shows the difference (number of frames) from the correct corresponding image.

られた。しかし、比較手法 2 と比較すると、正解率の向上は 3% であった。これは撮影条件が類似しており、得られる対応点の精度が高いことから、比較手法 2 においても安定した画像間距離を計算できたためである。Fig. 7 を見ると、比較手法 2 においても多くの画像で正しい結果が得られており、提案手法との差は大きくない。一方、日中に撮影した系列 5、早朝に撮影した系列 3 をそれぞれ入力とデータベースとした場合は、撮影条件の違いにより画像の見えが大きく異なる。このとき、提案手法は比較手法 1 に比べ 30%、比較手法 2 に比べ 28% 高い正解率であった。これは、撮影条件の違いから、入力とデータベース間で得られる対応点の精度が低く、そこから求める基礎行列の推定精度、そして画像間距離が不安定となったためである。そのため、本論文で提案する 1 対多画像間距離の効果が顕著に表れたと考えられる。Fig. 8 から、提案手法による検索精度向上が確認できる。

〈5・3〉 位置推定失敗とその傾向 Fig. 9 に、位置推定に失敗した例を示す。これは、天候が異なる系列 5 と系列 4 をそれぞれ入力とデータベースとした場合の内、約 100 フレームにわたって正しい画像の対応が得られず、その区間における提案手法の正解率が 19% まで低下したときの例である。ここでは、車体の大きな揺れに伴い、画像に大きなぶれが発生した。そのため、対応点を精度良く求めることができず、1 対 1 画像間距離および 1 対多画像間距離を安定して求めることができなかつた。このような位置推定の失敗は、対応点を求める際の精度向上により改善が可能で

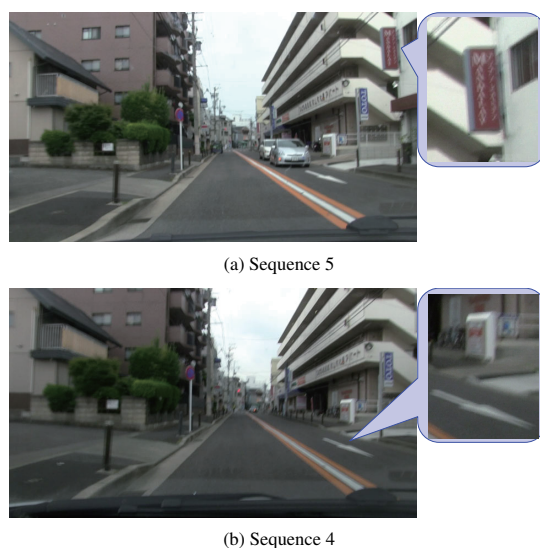


Fig. 9. Example of blurs caused by the vibration of the vehicle.

ある。まず第 1 に、撮影条件に合った適切な撮影パラメータでデータを取得することが 1 つの解決策となると考えられる。また近年、局所特徴量記述・対応点探索手法が盛んに研究されており、それらを利用することで、より高速・高精度に対応点を求めることができると考えられる<sup>(16)~(19)</sup>。これらに関しては、今後解決すべき課題である。

〈5・4〉 提案手法の適用条件 提案手法における 1 対多画像間距離 (式 (4)) は、複数フレームに対して 1 対 1 画像間距離  $1/|e_x|$  を求めることで計算される。もし、データ

ベースと完全に一致する走行経路で入力画像を取得した場合、1対1画像間距離の分母が $|e_x| = 0$ となり、式(4)を計算することができない。ただし、データベースと完全に一致する経路となることは極めて稀であるため、実用上は問題にならないと考えられる。また、走行経路の異なる2つの系列をデータベースとして用意し、自転車位置推定時にそれらを切り替えて用いることで、この問題を回避することも可能である。

提案手法では、1対多画像間距離が成立する条件として、式(4)に用いるデータベースの画像系列の撮影位置が等間隔であることを仮定している。しかし、常に等速で走行してデータを取得することは困難である。この問題は、画像系列の取得後に等間隔な撮影位置となるように再サンプリングしてデータベースを構築することで解決可能である。ただし本実験では、撮影位置を等間隔とする再サンプリングを行わず、撮影した映像に直接各手法を適用した。いずれも、可能な限り等速となるように走行して撮影した映像ではあるが、系列1, 4, 5については映像の終盤に赤信号による減速が含まれている。これらをデータベースとし、等速走行の系列3を入力とした場合、データベース系列の赤信号に伴う減速が生じた区間の正解率は85%であった。また、上記以外の等速で走行しているとみなせる区間の正解率は88%であった。これらより、データベースが等間隔に記録されていない場合、提案手法の精度は僅かに低下することがわかる。しかし、画像系列の撮影位置が等間隔となるよう再サンプリングしたデータベースを用いることでこの問題を回避可能であるため、実用上の大きな問題とならないと考えられる。

**〈5・5〉 計算時間** 本節では、提案手法の計算時間について考察する。GPSの位置情報を用いることで、データベースの検索範囲がだまかに絞り込まれていると仮定する。GPSの位置推定誤差を約10mとすると、本実験では、±30フレームの検索に相当する。この条件で提案手法を適用し、計算時間を評価した。CPU: Intel Core i7 Extreme Edition 990X, メモリ: 24 GBの計算機を用いたとき、入力1フレームあたりの計算時間は平均1.83秒であった。計算時間の内訳は、入力画像の特徴抽出に平均0.410秒、±30フレームの1対1画像間距離の算出に平均0.229秒、1対多画像間距離(式(4))を用いた画像検索に平均0.0050秒であった。近年、GPGPUやFPGAによる画像からの高速な特徴抽出手法が提案されており、入力画像からの特徴抽出は、数10から数100倍高速に計算可能である<sup>(20)~(22)</sup>。また、1対1画像間距離の算出は各フレーム対毎に並列化可能である。さらに、1対多画像間距離の式(4)を用いた画像検索は十分に高速である。これらのことから、提案手法による位置推定をリアルタイムで動作させることは十分可能であると考えられる。

一方、文献(9)の手法は正しい両端点の入力が必要なDPマッチングを用いているため、任意の1地点の位置を求めるために入力として画像系列を必要とする。そのため、系

列に含まれるフレーム数分の1対1画像間距離の算出が必要となる。これに対し、提案手法では入力となる1枚の画像に関する1対1画像間距離の算出のみを必要とする。このことから、提案手法は文献(9)の手法に比べ、計算コストの面で優位である。

## 6. まとめ

本論文では、市街地映像データベースが車両の進行方向に沿った系列として与えられる点に着目し、2カメラ間の幾何的な位置関係から導出される画像間距離を、データベースが時系列である点を考慮した新たな距離へと拡張した。そして、この距離に基づく市街地映像データベースからの画像検索により、車載カメラ画像1枚のみの入力から高精度に自転車位置推定可能な手法を提案した。実験から、提案手法は89%の正解率で対応画像を取得できることを確認した。これは、従来手法や単に1対1画像間距離が最小となる画像を求める手法に比べ高い正解率であり、提案手法の有効性を確認した。今後の課題として、高精度な対応点の算出手法の利用による提案手法の更なる精度向上や、DPマッチングとの組み合わせによる複合的な自転車位置推定手法の構築などが挙げられる。

## 謝辞

日頃より熱心に御討論頂く名古屋大学村瀬研究室諸氏に深く感謝する。本研究の一部は、JST 戦略的創造研究推進事業 CREST および科学研究費補助金による。また、本研究では画像処理にMISTライブラリ (<http://mist.murase.m.is.nagoya-u.ac.jp/>) を使用した。

## 文 献

- (1) 国土交通省:「カーナビ・VICS の出荷台数」, <http://www.mlit.go.jp/road/ITS/j-html/> (2013-2)
- (2) (一社) 情報通信ネットワーク産業協会:「プレスリリース “[2013年度 携帯電話の利用実態調査] を実施～スマートフォンが広く浸透。利用実態にも変化が…”」, <http://www.ciaj.or.jp/jp/pressrelease/pressrelease2013/2013/07/24/10785/> (2013-7)
- (3) (特非) ITS Japan:「通行実績・通行止情報」, <http://www.its-jp.org/saigai/> (2011-4)
- (4) 本田技研工業(株):「ニュースリリース “Honda インターナビの急ブレーキ多発地点データや交通事故情報などを活用した『SAFETY MAP』を一般公開”」, <http://www.honda.co.jp/news/2013/c130329a.html> (2013-3)
- (5) (一財) 日本自動車研究所:「移動体用高精度位置標定システムに関する調査研究報告書」, <http://www.jari.or.jp/tabid/347/Default.aspx> (2005-3)
- (6) J. Sato, T. Takahashi, I. Ide, and H. Murase: “Construction of a Street Image Map and Change Detection in Streetscapes from GPS Coordinated Omni-Directional Image Sequences”, IEICE Trans. Information and Systems, Vol. J90-D, No.4, pp.1085-1095 (2007-4) (in Japanese)  
佐藤准嗣・高橋友和・井手一郎・村瀬 洋:「GPS 座標付き全方位映像群からの市街地映像マップの構築と街並変化の検出」, 信学論(D), Vol. J90-D, No.4, pp.1085-1095 (2007-4)
- (7) S. Ono, R. Matsuhisa, H. Kawasaki, and K. Ikeuchi: “Matching Vehicle Camera Image Sequences with Spacetime Features”, Proc. Meeting on Image Recognition and Understanding 2010, pp.1145-1152 (2010-7) (in Japanese)  
小野晋太郎・松久亮太・川崎 洋・池内克史:「時空間特徴を用いた車載カメラ映像マッチング」, 画像の認識・理解シンポジウム 2010 講演予稿集, pp.1145-1152 (2010-7)
- (8) H. Uchiyama, D. Deguchi, T. Takahashi, I. Ide, and H. Murase: “Ego-Localization by Spatio-Temporal Matching between Videos Obtained from



Cameras with Different Viewing Angles Using an Extended DTW Method”, IEICE Trans. Information and Systems, Vol.J93-D, No.12, pp.2659–2665 (2010-12) (in Japanese)

内山寛之・出口大輔・高橋友和・井手一郎・村瀬 洋:「拡張 DP マッチングを用いた視野角の異なるカメラ映像間の時空間対応付けによる自車位置推定」, 信学論 (D), Vol.J93-D, No.12, pp.2659–2665 (2010-12)

- (9) H. Kyutoku, D. Deguchi, T. Takahashi, Y. Mekada, I. Ide, and H. Murase: “Ego-Localization Based on Frame Correspondences between In-Vehicle Camera Images and Streetview Database”, IEICE Trans. Information and Systems, Vol.J95-D, No.11, pp.1973–1982 (2012-11) (in Japanese)  
久徳遙矢・出口大輔・高橋友和・目加田慶人・井手一郎・村瀬 洋:「自車位置推定のための車載カメラ映像と市街地映像データベースの位置ずれや遮へいに頑健なフレーム対応付け」, 信学論 (D), Vol.J95-D, No.11, pp.1973–1982 (2012-11)
- (10) T. Yamasaki, T. Ishikawa, and K. Aizawa: “Retrieval of Images Captured by Car Cameras Using Its Front and Side Views and GPS Data”, IEICE Trans. Information and Systems, Vol.E90-D, No.1, pp.217–223 (2007-1)
- (11) D.G. Lowe: “Distinctive Image Features from Scale-Invariant Keypoints”, Int. J. Computer Vision, Vol.60, No.2, pp.91–110 (2004-11)
- (12) M. Susuki, T. Nakagawa, T. Sato, and N. Yokoya: “Extrinsic Camera Parameter Estimation from a Still Image Based on Feature Landmark Database”, Trans. VRSJ, Vol.13, No.2, pp.161–170 (2008-6) (in Japanese)  
薄 充孝・中川知香・佐藤智和・横矢直和:「ランドマークデータベースに基づく静止画像からのカメラ位置・姿勢推定」, 日本パッチャリリティ学会論文誌, Vol.13, No.2, pp.161–170 (2008-6)
- (13) S. Ueno, M. Hashimoto, A. Yoneyama, and R. Kawada: “A Study on Image Based Estimation of Position and Direction in Urban Environments Using SIFT Algorithm”, Tech. Rep. IEICE, PRMU2009-111 (2009-11) (in Japanese)  
上野智史・橋本真幸・米山暁夫・川田亮一:「SIFT を用いた都市環境の建物画像検索によるユーザ位置・方向推定手法に関する一検討」, 信学技報, PRMU2009-111 (2009-11)
- (14) 佐藤 淳:「コンピュータビジョン—視覚の幾何学—」, コロナ社 (1999)
- (15) M.A. Fischler and R.C. Bolles: “Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography”, Comm. ACM, Vol.24, No.6, pp.381–395 (1981-6)
- (16) Y. Ke and R. Sukthakar: “PCA-SIFT: A More Distinctive Representation for Local Image Descriptors”, Proc. 2004 IEEE Computer Society Conf. on Computer Vision and Pattern Recognition, Vol.2, pp.506–513 (2004-7)
- (17) T. Nishimura, S. Shimizu, and H. Fujiyoshi: “Keypoint Recognition Using Two-Stage Randomized Trees”, Proc. Meeting on Image Recognition and Understanding 2010, pp.1412–1419 (2010-7) (in Japanese)  
西村 孝・清水彰一・藤吉弘巨:「2 段階の Randomized Trees を用いたキーポイントの分類」, 画像の認識・理解シンポジウム 2010 講演予稿集, pp.1412–1419 (2010-7)
- (18) M. Ambai and Y. Yoshida: “CARD: Compact And Real-time Descriptors”, Proc. 13th IEEE Int. Conf. on Computer Vision, pp.97–104 (2011-11)
- (19) J. Heinly, E. Dunn, and J. Frahm: “Comparative Evaluation of Binary Features”, Proc. 12th European Conf. on Computer Vision, Vol.2, pp.759–773 (2012-10)
- (20) J. Svab, T. Krajinik, J. Faigl, and L. Preucil: “FPGA Based Speeded Up Robust Features”, Proc. 2nd IEEE Int. Conf. on Technologies for Practical Robot Applications, pp.35–41 (2009-11)
- (21) L. Yao, H. Feng, Y. Zhu, Z. Jiang, D. Zhao, and W. Feng: “An Architecture of Optimised SIFT Feature Detection for an FPGA Implementation of an Image Matcher”, Proc. 2009 IEEE Int. Conf. on Field-Programmable Technology, pp.30–37 (2009-12)
- (22) L. Chang, J. Hernandez-Palancar, L.E. Sucar, and M. Arias-Estrada: “FPGA-Based Detection of SIFT Interest Keypoints”, Machine Vision and Applications, Vol.24, No.2, pp.371–392 (2013-2)

久徳遙矢 (非会員) 2008 年名古屋大学工学部電気電子情報工学科卒業。2010 年同大学大学院情報科学研究科博士前期課程修了。修士 (情報科学)。現在、同研究科博士後期課程在学中。画像処理・パターン認識技術の基礎および ITS への応用に興味を持つ。電子情報通信学会学生員。



出口大輔 (非会員) 2003 年名古屋大学大学院情報科学研究科博士後期課程修了。博士 (情報科学)。2004～2006 年まで日本学術振興会特別研究員。2006 年名古屋大学大学院情報科学研究科研究員, 2006 年同大学大学院工学研究科研究員, 2008～2012 年まで同大学大学院情報科学研究科助教, 2012 年より同大学情報連携統括本部情報戦略室准教授。現在に至る。CARS2004 Poster Award, CADM2004 大会賞, 2006 年日本医用画像工学会奨励賞, 2006 年日本コンピュータ外科学会講演論文賞。電子情報通信学会, IEEE 各会員。



高橋友和 (非会員) 2003 年茨城大学大学院理工学研究科博士後期課程修了。博士 (工学)。同年より 2 年間名古屋大学大学院情報科学研究科 COE 研究員。2005 年より 3 年間日本学術振興会特別研究員。2008 年より岐阜聖徳学園大学経済情報学部准教授, 現在に至る。この間, 2013 年より 1 年間米



目加田慶人 (非会員) 1996 年名古屋大学大学院博士課程修了。1996 年宇都宮大学工学部情報工学科助手, 2001 年名古屋大学大学院工学研究科情報工学専攻助教, 2004 年中京大学生命システム工学部教授, 2008 年より同大学情報理工学部教授, 2013 年より同大学工学部教授。2006～2010 年名古屋大学大学院工学研究科特任教授を兼任。博士 (工学)。1998 年日本医用画像工学会論文賞, 2005 年 CADM 館野賞。電子情報通信学会シニア会員, IEEE, 日本医用画像工学会各会員。



井手一郎 (非会員) 2000 年東京大学大学院工学系研究科電気工学専攻博士課程修了。博士 (工学)。同年国立情報学研究所助手。2004 年名古屋大学大学院情報科学研究科助教, 2007 年より准教授。電子情報通信学会シニア会員, 情報処理学会, 映像情報メディア学会, 人工知能学会, 言語処理学会, IEEE, ACM 各会員。



村瀬 洋 (非会員) 1980 年名古屋大学大学院修士課程修了。同年日本電信電話公社 (現 NTT) 入社。1992 年から 1 年間米



国コロンビア大客員研究員。2003 年から名古屋大学大学院情報科学研究科教授, 現在に至る。博士 (工学)。1994 年 IEEE-CVPR 最優秀論文賞, 1995 年情報処理学会山下記念研究賞, 1996 年 IEEE-ICRA 最優秀ビデオ賞, 2001 年高柳記念奨励賞, 2002 年電子情報通信学会業績賞, 2003 年文部科学大臣賞, 2004 年 IEEE Trans. MM 論文賞, 2010 年前島密賞, 2012 年紫綬褒章, 他受賞。IEEE フェロー, 電子情報通信学会フェロー, 情報処理学会会員。