

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転車位置推定

内山 寛之[†] 出口 大輔[†] 高橋 友和^{††} 井手 一郎[†]
村瀬 洋[†]

Ego-Localization by Spatio-Temporal Matching between Videos Obtained from
Cameras with Different Viewing Angles Using an Extended DTW Method

Hiroyuki UCHIYAMA[†], Daisuke DEGUCHI[†], Tomokazu TAKAHASHI^{††}, Ichiro IDE[†],
and Hiroshi MURASE[†]

あらまし 本論文では、データベース作成時に全方位カメラ、位置推定時に通常カメラを使用するアピランススペースの自転車位置推定手法を提案する。従来のアピランススペースの自転車位置推定では、データベース作成時と位置推定時で同一種類のカメラが使用されることが多い。しかし、位置推定精度やカメラの設置位置の制約などの点で問題となることがある。提案手法では、通常カメラ映像のフレームに対応する全方位カメラ映像のフレームと、そのフレーム中の領域を探索する。しかし、市街地では類似する物体が多く存在し、単一フレーム照合では正確な照合が難しい。そのため、複数フレーム照合を用いることで精度の向上を図る必要がある。これを実現するため、フレームと領域を考慮した映像間の照合コストを定式化する。そして、動的計画法によりこれを最小化するためのアルゴリズムを導入し、このアルゴリズムを用いて映像全体で最適に対応づける。データベース内のフレームには位置情報が付加されており、これにより自転車位置を得る。実際に市街地で撮影した映像を用いた評価実験の結果、自転車位置推定誤差が 2 m 以下の割合が、領域の対応付けを行わない手法で 77%、提案手法で 90% であり、提案手法の有効性を確認した。

キーワード 自転車位置推定, DP マッチング, 全方位カメラ, ITS

1. ま え が き

高精度な自転車位置推定は、高機能なカーナビゲーションシステムを実現する上で、基本的な要素技術の一つである。運転者を目的地まで正確かつ確実に案内するためには、詳細な自転車位置を取得する必要がある。また、正確な位置が得られれば、一時停止すべき箇所などで運転者に警告を行う等の運転者支援も実現できる。しかし、現在のカーナビゲーションシステムで利用されている普及型 GPS には、10 ~ 20 m の誤差が含まれており [1]、上述の機能を実現するための十分な精度が得られない。一方、RTK-GPS などのより高精度な

GPS を用いることで、数 cm の精度で位置推定を行うことができる。しかし、建物等によって生じる衛星電波の遮へいやマルチパスに弱いという欠点がある。これらの問題を解決するために、車載カメラと GPS を組み合わせた自転車位置推定手法が検討されており、普及型 GPS の精度を上回る精度が得られることが報告されている [2]。

カメラを用いた自転車位置推定の一般的な枠組みは以下のとおりである。まず、RTK-GPS や慣性航法装置等の高精度な自転車位置推定用機材を搭載した専用車両を用い、あらかじめ市街地の画像とその地点の高精度な位置座標を収集して蓄積する。これにより、自転車位置推定のためのデータベースが構築される。構築するデータベースは、経路に沿って撮影した画像系列からなり、画像と正確な位置座標が組となって格納されている。自転車位置推定を行う際は、まず普及型 GPS 等で大まかな自転車位置を絞り込む。そして、車載カメラ

[†] 名古屋大学大学院情報科学研究科, 名古屋市
Graduate School of Information Science, Nagoya University,
Nagoya-shi, 464-8601 Japan

^{††} 岐阜聖徳学園大学経済情報学部, 岐阜市
Faculty of Economics and Information, Gifu Shotoku
Gakuen University, Gifu-shi, 500-8288 Japan

からの入力画像とデータベース中の画像を照合することで、それに付随する位置座標を得る。

自転車位置推定は、画像間を照合する方法によって、大きく二つに分類することができる。

- (1) 特徴点ベース [3] ~ [5]
- (2) アピアランスベース [2], [6] ~ [10]

特徴点を用いた手法は、あらかじめ特徴点をデータベースに格納しておき、入力画像から抽出した特徴点と照合を行う。そのため、画像そのものをデータベースに保持する場合と比べ、データ量を削減できるという利点がある。しかし、解像度の低い画像を用いる場合や、対象物がテクスチャの乏しい壁面や生い茂った樹木の場合など、環境や状況によっては特徴点を安定して抽出できないという問題がある。アピアランスベースの手法は、データベースに蓄積されている画像と入力画像を直接比較する。画像を直接比較するため、画像の解像度が低い場合や、特徴点の抽出が難しいシーンでも照合を行うことができる。また、データ量の問題は、近年の記憶装置の大容量化により解決が可能になってきている。

従来のアピアランスベースの自転車位置推定の手法では、データベース構築時と位置推定時に同一のカメラを用いる。例えば、通常カメラを搭載する構成 [6] ~ [8] や全方位カメラを搭載する構成 [2], [9], [10] が用いられる。通常カメラを用いた構成は、一般車両にも容易に搭載可能であるという利点がある。しかし、視野角が小さいため、カメラの装着位置の違いや、データベース構築時と位置推定時の走行軌跡の違いによる影響を大きく受ける。全方位カメラを用いた構成は視野角が大きいため、走行軌跡の違いによる影響に対して頑健であり、ロボットナビゲーションに広く用いられている。しかし、全方位カメラは広範囲を見渡すように設置する必要があるため、上部に高く張り出した構造にならざるを得ず、一般車両に搭載するには課題が多い。

そこで本論文では、データベース構築時に全方位カメラ、位置推定時に通常カメラを用いる手法を提案する。ここで提案するカメラ構成では、全方位カメラと通常カメラの視野角の違いにより、画像の見え方が大きく異なる。そのため、画像間の正確な照合が難しい。そこで本研究では、画像の見えの違いに対応すると同時に、画像系列の時間方向を対応づけることで、全方位カメラ映像と通常カメラ映像のフレーム間を対応づ

ける。また、市街地には類似する物体が多く存在するため、単一フレームの画像照合では正確な対応付けが難しい。そこで、複数フレームを照合に用い、時系列情報を利用することで、位置推定精度を向上させる。

以降、2. で提案手法の基本戦略を述べ、3. で本論文で提案する全方位カメラ映像と通常カメラ映像の対応付け手法を説明する。4. で評価実験とその考察について述べる。

2. 提案手法の基本戦略

ここでは全方位カメラ映像と通常カメラ映像の対応付けに従来手法を用いる場合の問題点と、提案手法の基本戦略について述べる。

2.1 従来手法を異種カメラ間対応付けに適用する場合の問題点

従来、時系列情報を利用する同一カメラ映像間を用いた自転車位置推定手法として、DP マッチングを用いた手法 [2], [6], [10] が提案されている。DP マッチングは、ある二つの時系列パターンに対し、パターンの順序関係を保存しつつ、時間軸の非線形な伸縮を許容する対応付けを求める手法である。これにより、車両の速度の違いに頑健な入力画像系列とデータベース画像系列間の対応付けが実現できる (図 1 (a))。佐藤ら [2] は、全方位カメラ映像同士の対応付けに DP マッチングを用いており、その原理は以下のとおりである。まず、データベース画像系列 $\{A[1], \dots, A[t], \dots, A[T]\}$ と入力画像系列 $\{B[1], \dots, B[u], \dots, B[U]\}$ が与えられるとする。データベース画像系列には位置情報が付加されており、入力画像系列の時刻 t とデータベース画像系列の時刻 u の対応を求めることにより、入力画像の撮影地点の位置を取得することを目的とする。文献 [2] では、画像間の非類似度を系列要素間の照合コスト $d(t, u)$ とする。DP マッチングは、 $d(t, u)$ の総和を系列全体の照合コストとし、それを最小とする (t, u) の軌跡を動的計画法によって計算する (図 2 (a))。その (t, u) の軌跡が、二つの画像系列の時刻間の対応となる。DP マッチングによる画像系列間の対応付けは、以下のように $d(t, u)$ の累積を評価式とし、それを最小化する t_1, t_2, \dots, t_U を決定する問題として定式化される。

$$F = \min_{t_1, t_2, \dots, t_U} \sum_{i=1}^U d(t_i, i) \quad (1)$$

$$d(t, u) = \|A[t] - B[u]\|. \quad (2)$$

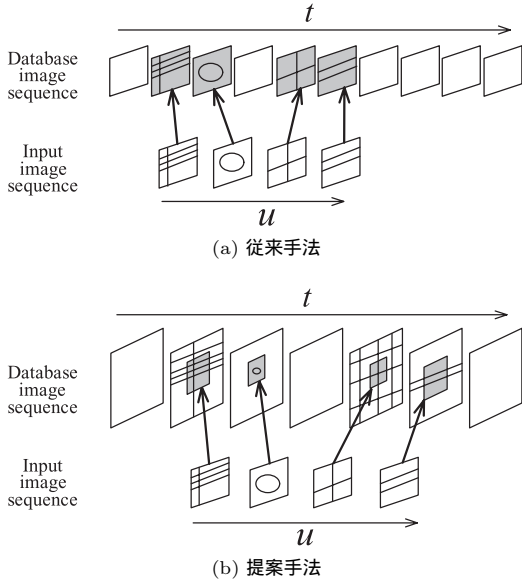


図 1 DP マッチングによる自転車位置推定．従来手法では時間軸方向のみ探索を行う．提案手法では，時間軸方向に加え，空間方向も探索を行う．

Fig. 1 Ego-localization using DP matching. The conventional method searches along the time direction only. The proposed method searches along the space direction in addition to the time direction.

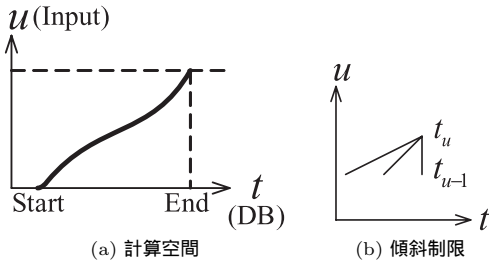


図 2 従来の DP マッチングの (a) 画像系列間の照合コストを最小とする対応付けの軌跡と (b) 傾斜制限の例
 Fig. 2 Conventional DP matching: (a) an optimal matching path between image sequences, (b) an example of a path constraint.

式 (2) の $\|a - b\|$ は a, b 間の距離を表す．ただし， t_1, t_2, \dots, t_U の選択の方法は，図 2 (b) に示すような傾斜制限による制約を受ける．

文献 [2] のように，同一カメラ映像間の照合を行う場合は，入力画像系列とデータベース画像系列の時間 t, u を対応づける．しかし，これを全方位カメラ映像と通常カメラ映像のフレーム間の対応付けに適用する場合，カメラ間の視野角の違いが問題となる．また，車両の位置ずれ，走行中の振動等によって生じる画像の

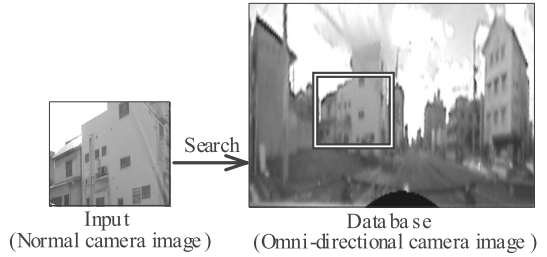


図 3 空間方向の探索．パノラマ展開された全方位カメラ画像上にサブウィンドウを設置し，入力画像との照合を行う．

Fig. 3 Search along the time direction. A sub-window is placed on a panoramic omnidirectional camera image to compare with an input image.

見えの違いも生じる．そのため，データベース画像系列の時間軸方向だけでなく，空間方向の対応部分も全方位カメラ映像中から探索する必要がある (図 3)．また，市街地では建築物の窓枠等，類似する部分が多く存在するため，正確な対応付けを行うためには，空間方向の探索に対しても時系列情報の利用が必要である．

2.2 基本戦略

提案手法では，全方位カメラ画像系列と通常カメラ画像系列の対応付けのために，次の二つのアプローチをとる．

- 全方位カメラ画像系列と通常カメラ画像系列の時間軸方向と空間方向の同時対応付け
- 時系列情報を利用した精度の向上

提案手法では，時間軸方向と空間方向の対応付けを定式化し，これを動的計画法で解くことで，図 1 (b) のように時間方向と空間方向の対応付けを同時に行う．これは，DP マッチングの拡張と考えられるため，ここでは拡張 DP マッチングと呼ぶ．全方位カメラは通常カメラに比べ視野角が大きいので，全方位カメラ映像にサブウィンドウを設置し，サブウィンドウと入力画像を照合する．そして，サブウィンドウの平行移動と拡大縮小を行うことで，見えの変動に対応する．次に，画像系列間の照合コストの評価式に，全方位カメラ映像上に設置したサブウィンドウの平行移動・拡大縮小を制御するパラメータを導入する．動的計画法を用いてこの評価式を最小化することで，系列全体として最適となる画像系列間の対応付けを求める．

3. 異種カメラ間の対応付け手法

提案手法では，全方位カメラ画像系列と通常カメラ

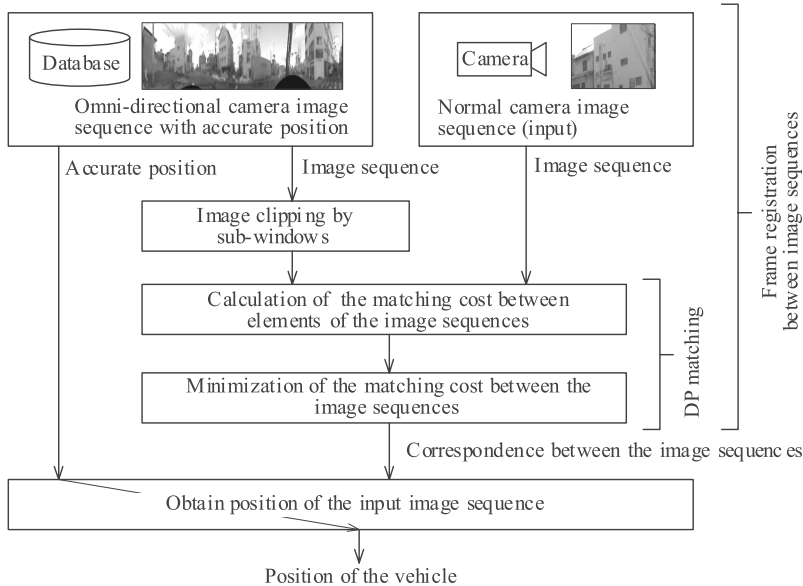


図 4 自車位置推定の流れ
Fig. 4 Flow of the ego-localization.

画像系列の間を拡張 DP マッチングにより対応づける。そのためには、動的計画法が適用できるように、画像系列間の照合コストを適切に設定する必要がある。具体的には、最小化する評価式を、動的計画法の漸化式に分解できるような形で表す必要がある。

自車位置推定の全体の流れを図 4 に示す。以降、全方位カメラ映像と通常カメラ映像のフレーム間対応付けの詳しい処理の内容を説明する。

3.1 画像間の照合

データベース構築時に全方位カメラを使用し、位置推定時に通常カメラを使用する場合、カメラの違いや撮影条件の違いにより、それぞれで得られる画像系列は見えが異なる。二つの画像系列間の差異は、次の二つに分けられる。

- 時間的な差異
- 空間的な差異

時間的な差異は、車両の走行速度の違いより発生する画像系列の時間方向の伸縮である。空間的な差異は、視野角の違いや視点の違いである。

まず、視野角の違いに対応するために、全方位カメラ上にサブウィンドウを設置し、サブウィンドウと入力画像を照合する。本論文では、全方位カメラと通常カメラを設置する位置は異なり、各カメラの水平方向は既知であるという状況を想定する。同一車両の同一

位置にカメラを設置し、かつ車両の走行軌跡が同じ場合は、視点の違いを考慮する必要はなく、サブウィンドウの位置や大きさを固定すればよい。しかし実際には、カメラを設置する位置の違いや車両の走行軌跡の違いのほか、走行中の振動で生じるカメラの角度変化により視点の違いが発生する。そこで、サブウィンドウの平行移動と拡大縮小を行うことで、見えの違いに対処する。車両が直進する場合には、サブウィンドウの水平方向の平行移動は変化が小さい。しかしながら、鉛直方向の平行移動は車両の走行軌跡の違いによって大きく変化する。そのため、平行移動は鉛直方向のみ行う。

まず、データベース画像系列 $\{A[1], \dots, A[t], \dots, A[T]\}$ と入力画像系列 $\{B[1], \dots, B[u], \dots, B[U]\}$ が与えられる。次に、時刻 t のデータベース画像から、平行移動・拡大縮小を制御するパラメータ v, s に対応するサブウィンドウを切り出す。以降、これを $\{A[v, s, t]\}$ と表記する。サブウィンドウ $A[v, s, t]$ と入力画像 $B[u]$ の照合コストは次式のように表される。

$$d(v, s, t, u) = \|A[v, s, t] - B[u]\|. \quad (3)$$

平行移動と拡大縮小を制御するパラメータ v, s に対し、サブウィンドウの平行移動の量と拡大縮小の倍率は一定値刻みで設定する。カメラ間の色合いの違い

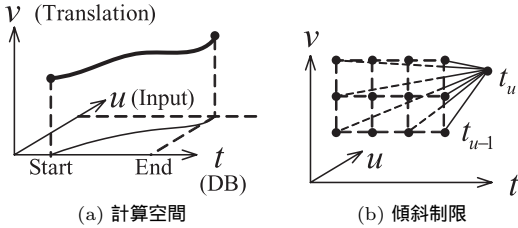


図 5 (a) 画像系列間対応付けコスト最小化処理のための計算空間．時刻 t, u に加え、サブウィンドウの平行移動を制御するパラメータ v を画像系列間の対応付けに導入する．(b) 極端な変化を抑制する制約を与える．

Fig. 5 (a) The calculation space for the optimization of the cost function. The proposed method introduces a sub-window parameter v in addition to time parameter t and u for matching the image sequences. (b) Matching path is constrained to avoid unusual changes.

を軽減するために、入力画像とデータベース画像の各サブウィンドウに色ヒストグラムの平坦化処理を行う。また、画像間の距離には、サブウィンドウの面積により正規化した画像の画素値からなるベクトル間の L1 距離を用いる。

3.2 画像系列間の照合コスト

データベース画像系列の時間軸方向のパラメータ t に加え、サブウィンドウの平行移動・拡大縮小を制御するパラメータ $v = 1, 2, \dots, V$ 及び $s = 1, 2, \dots, S$ を画像系列間の照合コストの評価式に導入する。自車位置推定を行うためには、ある入力画像系列の時刻に対応するデータベース画像系列の時刻を一意に決定する必要がある。すなわち、ある u に対応する v, s, t の組が一つだけ存在する必要がある。そのため、提案する拡張 DP マッチングは、図 5(a) のように、照合コストが系列全体で最小となるように、 u に対する v, s, t の軌跡を $\{(v_1, s_1, t_1, 1), \dots, (v_u, s_u, t_u, u), \dots, (v_U, s_U, t_U, U)\}$ として求める。図 5 では、拡大縮小を制御するパラメータ s を省略し、 $v-t-u$ の三次元の空間で表示しているが、実際は $v-s-t-u$ の四次元の空間である。

ある u に一つの v, s, t の組が対応するため、系列全体での照合コストを次式で表す。

$$F = \min_{(v_1, s_1, t_1)} \sum_{i=1}^U d(v_i, s_i, t_i, i) \quad (4)$$

$$\vdots$$

$$(v_U, s_U, t_U)$$

この式は、後の 3.4 で述べるように動的計画法で

解くことができる。動的計画法では、系列全体での照合コストが最小となるように $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ を決定する。また、式 (1) と同様に、 $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ の選択には、図 5(b) のような傾斜制限により制約を与える。

3.3 傾斜制限

画像系列間の誤対応を抑制するため、映像の撮影条件に応じ、式 (4) の $(v_1, s_1, t_1), \dots, (v_U, s_U, t_U)$ の選択方法に制約を加える。提案手法では、次のような仮定を設ける。

(a) 車両の最大速度は限られており、自車位置が短時間に大きく変化することはない

(b) 車両は前進するのみで後退しない

(a) と (b) の仮定により、入力画像系列の時刻 $u-1$ に対応するデータベース画像系列の時刻 t_{u-1} と、時刻 u に対応する時刻 t_u の差はしきい値 θ_t 以内とする。

$$0 \leq t_u - t_{u-1} \leq \theta_t \quad (5)$$

次に、(a) の仮定により、平行移動・拡大縮小の変化量を制限できる。例えば、時刻 u に対応する平行移動のパラメータ v_u は、時刻 $u-1$ に対応する v_{u-1} の近傍になる。そのため、平行移動の条件はしきい値 θ_v を用いて次式のように表される。

$$-\theta_v \leq v_u - v_{u-1} \leq \theta_v \quad (6)$$

同様に、拡大縮小の条件もしきい値 θ_s を用いて次式のように表される。

$$-\theta_s \leq s_u - s_{u-1} \leq \theta_s \quad (7)$$

3.4 拡張 DP マッチングのアルゴリズム

系列間の照合コストが最小となるように、式 (4) を $\{(v_1, s_1, t_1), \dots, (v_u, s_u, t_u), \dots, (v_U, s_U, t_U)\}$ について最小化する。その際、照合が系列全体で最適となるように、動的計画法 [11] を用いて解く。

ここで、拡張 DP マッチングのアルゴリズムを導出するために、入力画像系列の部分系列 ($u = 1, 2, \dots, U'$) とデータベース画像系列を照合する問題を考える。この照合コストを $g(\mathbf{a}_{U'}, U')$ とおく。ただし、簡略化のため、 $\mathbf{a} = (v, s, t)$ と表記する。これを次のように変形する。

$$g(\mathbf{a}_{U'}, U') = \min_{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{U'}} \sum_{i=1}^{U'} d(\mathbf{a}_i, i)$$

$$= \min_{\mathbf{a}_{U'}} \left[\min_{\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_{U'-1}} \sum_{i=1}^{U'-1} d(\mathbf{a}_i, i) + d(\mathbf{a}_{U'}, U') \right]$$

よって、次式が求まる。

$$g(\mathbf{a}_{U'}, U') = \min_{\mathbf{a}_{U'}} [g(\mathbf{a}_{U'-1}, U' - 1) + d(\mathbf{a}_{U'}, U')]$$

これは、入力画像系列の部分系列 ($u = 1, 2, \dots, U'$) とデータベース画像系列の照合コストを最小にするためには、更にその部分系列 ($u = 1, 2, \dots, U' - 1$) の照合コストを最小化したものを用いばよいことを示す。その結果、式 (4) は次の漸化式を $u = 1, 2, \dots, U$ について計算することで最小化を行うことができる。

$$g(v, s, t, u) = \min_{\substack{-\theta_v \leq p_v \leq \theta_v \\ -\theta_s \leq p_s \leq \theta_s \\ 0 \leq p_t \leq \theta_t}} \left\{ g(v - p_v, s - p_s, t - p_t, u - 1) + d(v, s, t, u) \right\} \quad (8)$$

図 5 にこの漸化式の計算空間の概略を示す。このとき、式 (4) は次式として求まる。

$$F = \min_{v, s, t} g(v, s, t, U) \quad (9)$$

式 (9) で最小値を与える v, s, t が、入力画像 $u = U$ に対応するサブウィンドウの制御パラメータ及びデータベース画像となる。このときのデータベース画像系列の時刻 t に対応する位置が現在の推定位置となる。具体的な拡張 DP マッチングのアルゴリズムを付録に示す。

大規模なデータベースを用いる場合、 t のとり得るすべての範囲に対して式 (8) を計算することは、記憶容量・計算量の観点から現実的ではない。そこで、自車位置推定の開始時に人手または普及型 GPS 等で大まかな位置を絞り込み、整合窓 [12] を導入することで、計算範囲を限定する。

4. 実験

4.1 実験条件

提案手法の有効性を確認するために、位置推定誤差を以下の実験により評価した。実験に用いるカメラの車両への装着位置を図 6 に示す。データベース構築には円錐ミラー型の全方位カメラを使用し、路面から 1.5 m の位置に設置した。自車位置推定用の通常カメラは民生用の DV カメラを用い、左前方へ向け、車内



Omni-directional camera for constructing a database
Normal camera for ego-localization

図 6 カメラを設置した位置
Fig. 6 Camera positions.

の助手席のサイドガラスに固定した。また、通常カメラは路面から 1.0 m の位置に設置し、設置方向は人手で測定した。撮影した映像のフレームレートは全方位カメラ、通常カメラともに 30 fps であった。

経路長が 300 m と 400 m の二つの区間を用い、合計で 7 回の実験を行った。映像中には車両の加減速、停止が含まれている。なお、データベース構築用と自車位置推定用の映像は異なる日付に撮影したものを使用した。映像収集期間は約 6 か月である。

データベース車両と位置推定車両の速度の差が極端に大きい場合、コスト最小化の処理の際に問題となる。そこで、データベース内の画像系列の 1 フレーム当りの走行距離が、0.4 m となるように正規化したデータベース画像系列を用意した。

提案手法におけるパラメータは予備実験により得られた以下の値を使用した。まず、式 (5) の θ_t は 3 とした。これにより、位置推定車両の速度は 0 ~ 130 km/h に対応可能である。画像間の照合における画像の平行移動は 1.44 度刻み、拡大縮小は 0.03 倍刻みとした。また、 θ_v と θ_s は 1 とした。カメラの内部パラメータはカタログ値を使用した。入力画像の下半分は道路領域のため、上半分のみを画像照合に使用した。自車位置推定開始時の初期探索範囲として、真の自車位置を中心とした約 120 m の区間を人手で与えた。全方位カメラ映像の元の解像度は 640×480 pixel であった。パノラマ展開は球面投影によって行い、水平方向の解像度は 180 度当り 250 pixel とした。また、入力画像の元の解像度は 640×480 pixel であり、パノラマ展開後の全方位カメラの解像度と通常カメラの画角を考慮し、拡大縮小の倍率が 1 のときの解像度を 61×56 pixel として位置推定を行った。なお、パノラマ展開の際に、全方位カメラと通常カメラのカメラ中心は同位置にあると仮定した。

本実験では、表 1 に示すように、提案手法を二つの比較手法と比較した。比較手法 1 は画像の平行移動・拡大縮小を行わず、式 (1) の従来の DP マッチングを用いた。すなわち、画像間の距離を算出する際に、サ

表 1 実験条件
Table 1 Experimental condition.

| | 平行移動・拡大縮小 | DP マッチング |
|--------|-----------|----------|
| 比較手法 1 | なし | 従来 |
| 比較手法 2 | あり | 従来 |
| 提案手法 | あり | 提案 |

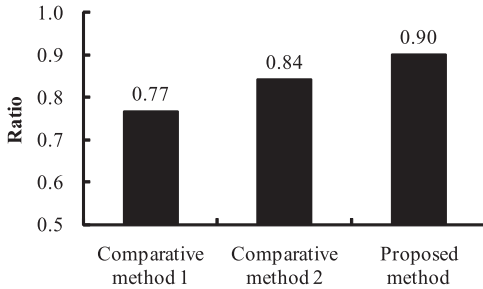


図 7 位置推定誤差が 2.0m 以内に収まる割合

Fig. 7 The ratio which the ego-localization error is less than 2.0m.

ブィンドウを一定位置かつ一定サイズに固定し、系列要素間の照合コスト $d(t, u)$ を算出した。比較手法 2 は画像の平行移動・拡大縮小を行い、従来の DP マッチングを用いた。具体的には、フレーム単位でサブウィンドウの平行移動と拡大縮小を総当りで行い、最小となる照合コストを系列要素間の照合コスト $d(t, u)$ とした。提案手法はサブウィンドウの平行移動・拡大縮小を行い、提案する拡張 DP マッチングを用いた。

実験結果の数値はフレーム単位で算出した。ただし、車両が停止している間は推定位置の変化が生じないため、車両が停止しているフレームを除いて評価した。

4.2 結果と考察

進行方向の誤差が 2m 以内に収まる割合を図 7 に示す。また、提案手法と比較手法の位置推定誤差の累積分布を図 8 に示す。図 7、図 8 より、提案手法、比較手法 2、比較手法 1 の順で良い結果が得られることが分かる。

比較手法 1 は、カメラの装着位置の違いや走行位置の違い、カメラの振動により発生する入力画像の見た目の変化を考慮しておらず、提案手法や比較手法 2 に比べ、位置推定精度が低かった(図 7、図 8)。提案手法が成功し、比較手法 1 が失敗した例を図 9 に示す。(b) と (c) 中の実線の矩形は (a) の入力画像に対して対応すると推定されたサブウィンドウである。(c) 中の点線の方形は正しいサブウィンドウの位置を示している。提案手法では、(b) のように入力画像とデータ

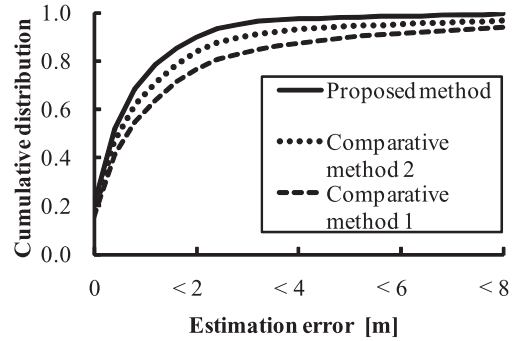
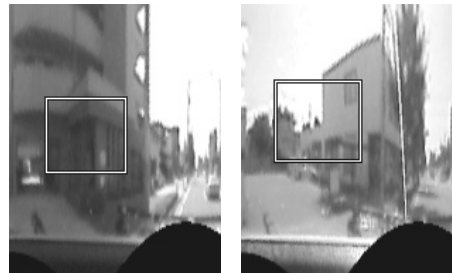


図 8 位置推定誤差の累積分布

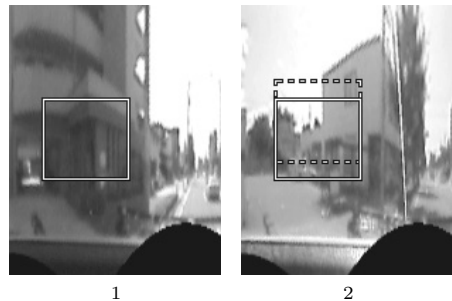
Fig. 8 The cumulative distribution of the ego-localization error.



(a) 入力画像系列



(b) データベース画像系列 (提案手法)



(c) データベース画像系列 (比較手法 1)

図 9 画像間の対応付け結果 (比較手法 1 との比較)

Fig. 9 Results of the matching between the image sequenes. Comparison between proposed method and comparative method 1.

ベース画像が正しく対応づいている。一方、比較手法 1 では、(c) の 1 番目の画像のように入力画像とデータベース画像が一致する場合があるものの、2 番目の

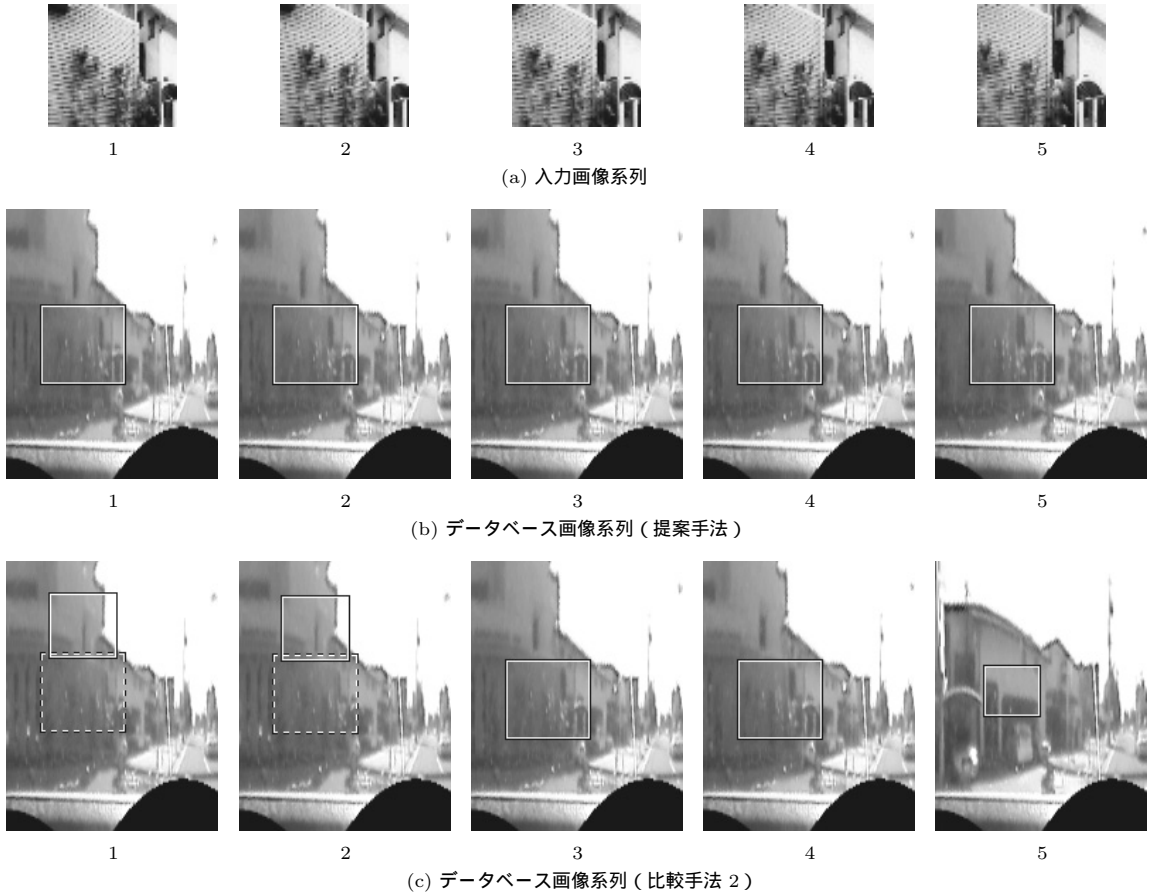


図 10 画像間の対応付け結果 (比較手法 2 との比較)

Fig. 10 Results of the matching between the image sequences. Comparison between the proposed method and comparative method 2.

画像のように画像系列間の対応付けを誤る場合が多く見られた．このように，提案手法と比較手法 1 との比較から，全方位カメラ映像と通常カメラ映像間の照合のためには，画像の空間方向の探索も行う必要があることが確認できる．

次に，提案手法が成功し，比較手法 2 が失敗した例を図 10 に示す．(a) は入力画像系列の連続する 5 フレームであり，(b)，(c) はそれに対応すると推定されたデータベース映像中のフレームを示している．(b) と (c) 中の実線の矩形は (a) の入力画像に対して対応すると推定されたサブウィンドウであり，点線の方形は正しいサブウィンドウのおおよその位置を示している．提案手法では (b) のように，サブウィンドウは入力画像と同じ部分を示している．このように，入力画像系列に対応するデータベース画像系列のフレームと，

サブウィンドウの位置や大きさの推定に成功していることが分かる．一方，比較手法 2 では，フレーム単位でサブウィンドウの平行移動と拡大縮小を総当りで行う．そのため，データベース画像上に類似する部分が複数存在する場合には，(c) の 1, 2, 5 番目の画像のように画像系列間の対応付けを誤ることがあった．(c) の 1, 2 番目の画像は，推定されたフレームはほぼ正しいものの，サブウィンドウの位置や大きさの推定に失敗した．5 番目の画像は，サブウィンドウの位置や大きさの推定に加え，フレームの推定も失敗した．このように，提案手法と比較手法 2 の比較から，画像系列の時間軸方向だけではなく，空間方向の時系列情報も考慮する提案手法が有効であることが確認できる．

入力画像にデータベース画像には存在しない物体が映った場合には，自転車位置推定の失敗が見られた．例

えば、図 11 (a), (b) のような車高の高い路上駐車車両が入力画像領域の大部分を遮へいしている間は位置推定誤差が増加した。しかしながら、遮へいされた時間はいずれも 1 秒程度であり、遮へいが終了した直後には、正しい自車位置推定を再開した。このように、駐車車両による遮へいは短時間であり、影響は少ないと考えられる。一方で、図 11 (c) のような車高の低い車両や歩行者は、画像の一部分を遮へいするものの、画像中に占める割合が小さいため、影響は見られなかった。

実験に用いた映像には、晴天と曇天に撮影したものが含まれており、撮影時刻はいずれも 10:00 ~ 13:00 であった。この条件では、位置推定失敗の原因となるような建築物の陰影等の強い影響は見られなかった。しかし、今後より多様な光源環境の下で実験を行う必要がある。そのほかに、通常カメラを車内に設置したため、天候や太陽の方向によっては、ガラスへの映り込みが発生し、位置推定を誤る例が見られた。また、全方位カメラの特性上、強い直射日光により、画像の一部分の明るさが大きく変化する例も見られた。そのため、今後カメラの設置方法等を検討する必要がある。

そのほかに、提案手法において、図 12 のように距

離の異なる二つの建築物が同時に入力カメラに存在する際、位置推定誤差が増加する場合があった。これは、単純な平行移動と拡大縮小で、画像の見えの変化を近似し、単純な画素値に基づく画像特徴を用いたため、見えの変化への対処が不十分なためだと考えられる。画像照合の失敗により位置推定誤差が増加する場合には、推定される速度が急激に変化する。そのため、それを検知することで、位置推定精度が低下する前の速度情報から、推定位置を補間できるものと考えられる。

実験では、通常カメラは車両の左前方に向けたが、提案手法はその他の方向も適用可能である。そこで、通常カメラを左後方に向け、前述の実験と同じ経路で、計 7 回実験を行った。その結果、位置推定誤差が 2 m 以下の割合は 90% であった。左前方の場合も 90% であり、取付け位置による顕著な違いは見られなかった。

4.3 カメラの水平方向の変化に関する考察

提案手法では、車両が直進する場合を想定し、カメラを固定して設置した方向は既知とした。しかし、実際には、車線変更を行う等、車両が斜めに進む場合は、直進する場合に対して、カメラの水平方向が変化する。そのため、この影響を調査するために、全方位画像上で探索する方向として、真のカメラ方向とは異なる角度を設定し、自車位置推定実験を行うことで、カメラ方向の変化による影響を調べた。

真のカメラ方向と全方位画像上で探索する方向の間の角度と、位置推定誤差が 2.0 m 以内に収まる割合の関係を示したグラフを図 13 に示す。仮に車線幅 3 m の道路において、時速 60 km で走行中に 5 秒間で車線変更を行った場合、道路方向に対する傾きは平均約 2

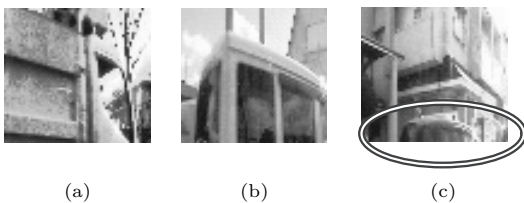


図 11 駐車車両による遮へい
Fig. 11 Occlusion caused by a vehicle.

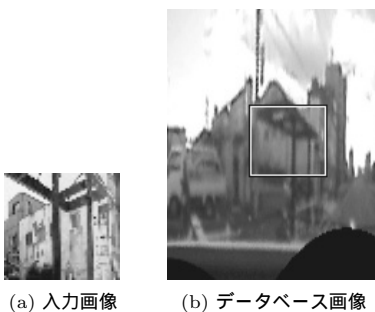


図 12 距離の異なる二つの建築物が同時に画像中に存在する場合、位置推定誤差が増加する
Fig. 12 The Estimation error increases when two structures exist in different distances.

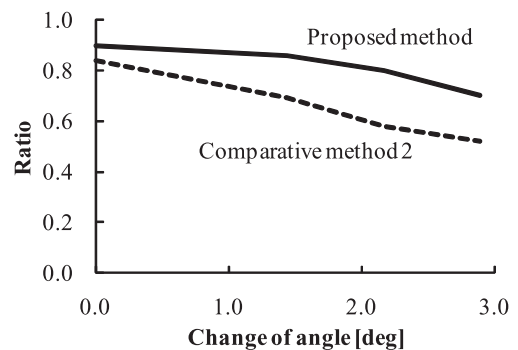


図 13 カメラの方向の変化と、位置推定誤差が 2.0 m 以内に収まる割合の関係
Fig. 13 The relationship between the change of the camera angle and the ratio which the ego-localization error is less than 2.0 m.

度である。この場合、図 13 より、提案手法の位置推定誤差が 2.0m 以内に収まる割合の低下は 1 割程度であった。車線変更は頻度が少なく、短時間であるため、カメラ方向の変化の影響は十分に小さいと考えられる。また、角度のずれが大きい場合でも、提案手法は比較手法 2 に比べ、位置推定精度の低下割合が少なかった。

4.4 実用上の課題に関する考察

実利用を考えた場合には、自転車位置推定の開始時に、自転車の初期位置が必要となる。およそその初期位置を得るために、現在広く使われている普及型 GPS が利用可能である。しかし、普及型 GPS には 10~20m の誤差が含まれるため、その誤差を考慮した範囲を参照画像系列として与え、提案手法により自転車位置を絞り込む。このようにすることで、精度の低い初期値を基に、高精度な自転車位置推定を実現する。

位置推定の推定頻度が少ない場合、測定ごとの位置推定精度が高くても、時間的な分解能が低下し、結果として位置推定精度が低下する。現在、広く自転車位置推定に用いられている普及型 GPS の時間分解能は 1 秒間に 1 回程度である。一方で、カメラの時間分解能 (フレームレート) は 30 fps である。そのため、カメラを用いることで高い時間分解能で自転車位置推定が実現できる。

実験環境は、CPU Xeon E5540 2.53 GHz × 2、メモリ 32 G バイトであり、処理時間は入力画像 1 フレーム当たり平均 0.78 秒であった。現状ではリアルタイム処理は困難であるが、計算機の性能向上やより高速な DP マッチング手法の導入により、処理速度の向上が期待できる。

また、本論文では直進のみを想定したが、実際には交差点等による右左折も考慮する必要がある。その場合には、交差する道路間で、使用する参照画像系列を切り換えることが必要となる。右左折を行う場合には車両の角度は大きく変化する。そのため、提案手法では水平方向のカメラ方向は固定としたが、水平方向の探索も導入する必要があると考えられる。また、参照画像系列の切替の際には、入力画像系列とそれらの参照画像系列との照合を行い、照合コストが小さい方の参照画像系列に付随する位置情報を採用するという方法が考えられる。

データベースの容量は、1 フレームの画像サイズを 4k バイト、フレームの撮影間隔を 0.4 m とすると、1 m 当り 10k バイトとなる。日本の道路の総延長が約 119 万 km であることから [13]、仮に日本すべての道

路をデータベース化すれば約 12 T バイト、1 都道府県当りにすれば平均約 253 G バイトである。実用上は、高精度な自転車位置が必要となる都市部等、必要な箇所だけをデータベース化すればよいため、更に少ない容量でよい。近年の記憶装置や通信技術の発達を考慮すれば、これは実用的な容量であると考えられる。

5. む す び

アピランススペースの手法による、自転車位置推定のための全方位カメラ映像と通常カメラ映像のフレーム間の対応付け手法を提案した。まず、全方位カメラ画像上にサブウィンドウを設置し、サブウィンドウと入力画像を比較した。そして、サブウィンドウの平行移動と拡大縮小を行うことで、カメラ間の見えの違いに対処した。また、平行移動と拡大縮小を制御するパラメータを拡張 DP マッチングに導入することで、画像系列の時間軸方向に加え、空間方向の時系列情報を利用した。実験の結果、空間方向の時系列情報を利用しない場合に比べ、位置推定精度が向上することを確認した。

今後の課題として、視点位置の違いに頑健な画像間の照合方法及び特徴量の検討、複数台の入力カメラを用いることによる位置推定精度向上等が挙げられる。また、本論文では、全方位カメラ映像と通常カメラ映像の対応付け手法に焦点を当てたが、提案手法の手順は広角の通常カメラと狭角の通常カメラ等の組合せにも適用が可能である。今後、このようなカメラ構成での実験を行う予定である。

謝辞 日ごろ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 及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] 白井澄夫, “高度衛星測位技術の動向と応用,” 信学誌, vol.92, no.9, pp.768–774, Sept. 2009.
- [2] 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, April 2007.
- [3] E. Royer, J. Bom, M. Dhome, B. Thuillot, M. Lhuillier, and F. Marmouton, “Outdoor autonomous navigation using monocular vision,” Proc. Int. Conf. on Intelligent Robots and Systems (IROS) 2005, pp.3395–3400, Aug. 2005.

- [4] 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” 日本バーチャルリアリティ学論, vol.13, no.2, pp.161-170, June 2008.
- [5] N. Yazawa, H. Uchiyama, and H. Saito, “Image based view localization system retrieving from a panorama database by SURF,” Proc. IAPR Conf. on Machine Vision Applications (MVA) 2009, pp.118-121, May 2009.
- [6] 小島 浩, 伊藤慶明, 岡 隆一, “Reference Interval-Free 連続 DP を利用した移動ロボットの時系列画像による位置同定システム” 信学論 (D-II), vol.J80-D-II, no.3, pp.724-733, March 1997.
- [7] 石川敬之, 山崎俊彦, 相澤清晴, “GPS と画像特徴を用いた車載ライフログ映像処理” 画像の認識・理解シンポジウム (MIRU2005) 論文集, pp.604-696, July 2005.
- [8] 前田佐嘉志, 久野義徳, 白井良明, “固有空間解析に基づく移動ロボットの位置認識” 信学論 (D-II), vol.J80-D-II, no.6, pp.1502-1511, June 1997.
- [9] 岩佐英彦, 粟飯原述宏, 横矢直和, 竹村治雄, “全方位画像を用いた記憶に基づく位置推定” 信学論 (D-II), vol.J84-D-II, no.2, pp.310-320, Feb. 2001.
- [10] 小島 浩, 高橋勝彦, 関 進, 岡 隆一, “走行ロボット動画像からのモーションシーンのスポッティング認識” 信学技報, PRU93-58, Sept. 1993.
- [11] R. ベルマン, S. ドレイファス, 応用ダイナミック・プログラミング, 日本科学技術連盟, 1962.
- [12] 内田誠一, “DP マッチング概説—基本と様々な拡張” 信学技報, PRMU2006-166, Dec. 2006.
- [13] 国土交通省道路局, “道路統計年報 2007” Nov. 2007.

付 録

拡張 DP マッチングのアルゴリズムを以下に示す.

[Step 1] Initialization

計算空間の範囲外に ∞ を代入する. すなわち, 範囲が $1 \leq v \leq V$ かつ $1 \leq s \leq S$ かつ $1 \leq t \leq T$ かつ $1 \leq u \leq U$ 外するとき,

$$g(v, s, t, u) = \infty \quad (\text{A}\cdot 1)$$

[Step 2] Iteration

$u = 1, 2, \dots, U$ について, 次の漸化式を計算する.
for $u = 1$

for $v = 1, 2, \dots, V, s = 1, 2, \dots, S, t = 1, 2, \dots, T$

$$g(v, s, t, 1) = d(v, s, t, 1) \quad (\text{A}\cdot 2)$$

for $u = 2, 3, \dots, U$

for $v = 1, 2, \dots, V, s = 1, 2, \dots, S, t = 1, 2, \dots, T$

式 (8) を計算する.

[Step 3] Output

入力画像 U が入力されたときに, 画像系列間の照

合コストが最小となるパラメータ v_U, s_U, t_U を出力する. また, v_u, s_u, t_u ($u = 1, 2, \dots, U - 1$) が必要ならば, バックトラックを行う.

$$(v_U, s_U, t_U) = \arg \min_{(v, s, t)} g(v, s, t, U) \quad (\text{A}\cdot 3)$$

(平成 22 年 3 月 25 日受付, 7 月 25 日再受付)



内山 寛之 (学生員)

平 19 名工大・工・電気情報卒. 平 21 名大大学院情報科学研究科博士前期課程了. 現在, 同研究科博士後期課程在学中. 画像処理・パターン認識の分野に興味をもつ. 平 22 電気関係学会東海支部大会 IEEE 名古屋支部学生奨励賞受賞.



出口 大輔 (正員)

平 13 名大・工・情報卒. 平 18 同大大学院博士後期課程了. 平 16~18 日本学術振興会特別研究員. 平 18 名大大学院情報科学研究科研究員, 平 18 名大大学院工学研究科研究員, 平 20 より同大助教, 現在に至る. 主に画像処理・パターン認識技術の開発とその ITS 及び医用応用に関する研究に従事. 博士 (情報科学). CARS2004 Poster Award, CADM2004 大会賞, 平 18 日本医用画像工学会奨励賞, 平 18 日本コンピュータ外科学会講演論文賞受賞. 日本医用画像工学会, 日本コンピュータ外科学会, 日本生体医工学会各会員.



高橋 友和 (正員)

平 9 茨城大・工・情報卒. 平 12 同大大学院理工学研究科博士前期課程了. 平 15 同研究科博士後期課程了. 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. 平 17 より 3 年間日本学術振興会特別研究員. 平 20 より岐阜聖徳学園大学経済情報学部准教授, 現在に至る. 画像認識の基礎研究並びにその応用に興味をもつ. 画像電子学会会員.



井手 一郎 (正員:シニア会員)

平 6 東大・工・電子卒。平 8 同大大学院工学系研究科情報工学専攻修士課程了。平 12 同研究科電気工学専攻博士課程了。博士(工学)。同年国立情報学研究所助手。平 16 名古屋大学大学院情報科学研究科助教授, 情報・システム研究機構国立情報学研究所客員助教授兼任。平 19 より准教授。この間, 平 14~16 総合研究大学院大学数物科学研究科助手併任, 平 17, 18, 19 フランス情報学・統計システム研究所(IRISA) 招聘教授。平 22~23 オランダアムステルダム大学情報学研究所上級訪問研究員。パターン認識技術の実応用や映像メディア処理全般に興味をもっている。情報処理学会, 映像情報メディア学会, 人工知能学会, IEEE Computer Society, ACM 各会員。



村瀬 洋 (正員:フェロー)

昭 53 名大・工・電気卒。昭 55 同大大学院修士課程了。同年日本電信電話公社(現 NTT) 入社。平 4 から 1 年間米国コロンビア大客員研究員。平 15 から名古屋大学大学院情報科学研究科教授, 現在に至る。文字・図形認識, コンピュータビジョン, マルチメディア認識の研究に従事。工博。昭 60 本会学術奨励賞, 平 6 IEEE-CVPR 最優秀論文賞, 平 7 情報処理学会山下記念研究賞, 平 8 IEEE-ICRA 最優秀ビデオ賞, 平 13 高柳記念奨励賞, 平 13 本会ソサイエティ論文賞, 平 14 本会業績賞, 平 15 文部科学大臣賞, 平 16 IEEE Trans. MM 論文賞, 他受賞。IEEE フェロー, 情報処理学会会員。